UNIVERSITY
OF
TORONTO
LIBRARY

MICROFORMED BY
PRESERVATION
SERVICES
DATE. MAR. 11 18

REPERTORIUM
DER HÖHEREN ANALYSIS

UNTER MITWIRKUNG DER HERREN
R. FRICKE IN BRAUNSCHWEIG · PH. FURTWÄNGLER IN AACHEN · A. GULDBERG IN CHRISTIANIA · H. HAHN IN CZERNOWITZ · E. JAHNKE IN BERLIN
H. JUNG IN HAMBURG · A. LOEWY IN FREIBURG · E. PASCAL IN NEAPEL
H. E. TIMERDING IN BRAUNSCHWEIG

HERAUSGEGEBEN VON

PAUL EPSTEIN

PROFESSOR AN DER UNIVERSITÄT STRASSBURG I. E.

ZWEITE AUFLAGE

ERSTE HÄLFTE

ALGEBRA, DIFFERENTIAL- UND INTEGRALRECHNUNG

200914
26/2/26

LEIPZIG UND BERLIN
DRUCK UND VERLAG VON B. G. TEUBNER
1910

MICROFORMED BY
PRESERVATION
SERVICES
DATE.....

QA
37
P375
1910
Bd 1
T 1

COPYRIGHT 1910 BY B. G. TEUBNER IN LEIPZIG.

ALLE RECHTE,
EINSCHLIESSLICH DES ÜBERSETZUNGSRECHTS, VORBEHALTEN.

Germany

E. PASCAL
O. PROFESSOR AN DER KGL. UNIVERSITÄT [illegible]

REPERTORIUM DER HÖHEREN MATHEMATIK

ZWEITE, VÖLLIG UMGEARBEITETE AUFLAGE DER DEUTSCHEN AUSGABE, UNTER MITWIRKUNG ZAHLREICHER MATHEMATIKER

HERAUSGEGEBEN VON

P. EPSTEIN UND H. E. TIMERDING
IN STRASSBURG IN BRAUNSCHWEIG

ERSTER BAND
ANALYSIS

LEIPZIG UND BERLIN
DRUCK UND VERLAG VON B. G. TEUBNER
1910

VORREDE DES VERFASSERS DER ITALIENISCHEN ORIGINALAUSGABE.

„Gerade bei einem Buch, wie demjenigen, welches ich hiermit dem mathematischen Publikum vorlege, scheint es notwendig zu sein, vor allen Dingen über die Absicht, in welcher es geschrieben ist, Aufklärung zu geben, damit Mißverständnisse unter allen Umständen ausgeschlossen sind."

„Das Buch hat den Zweck, auf einem möglichst kleinen Raum die wichtigsten Theorien der neueren Mathematik zu vereinigen, von jeder Theorie nur so viel zu bringen, als nötig ist, damit der Leser sich in ihr orientieren könne, und auf die Bücher zu verweisen, in welchen er Ausführlicheres finden kann."

„Es soll für den Studierenden, welcher während seiner Universitätszeit sich mit verschiedenen Zweigen der Mathematik beschäftigt hat, ein 'Vademecum', ein Taschenbuch sein, in welchem er, kurz zusammengefaßt, alle jene mathematischen Begriffe und Resultate wiederfindet, die er während seiner Studien sich nach und nach angeeignet hat oder doch hätte aneignen sollen. Man würde sich daher irren, wenn man der Ansicht wäre, ich hätte eine Enzyklopädie der Mathematik schreiben wollen; für eine solche Arbeit würden weder meine Kräfte ausgereicht haben, noch hätte der verhältnismäßig geringe Umfang dieses Buches genügt. Ich habe weiter nichts als ein bescheidenes *Repertorium* abfassen wollen, welches, wie ich glaube, den Studierenden der Mathematik Dienste zu leisten imstande ist."

„Das Buch kann den jungen Mathematikern auch insofern von großem Nutzen sein, als es ihnen die Möglichkeit bietet, ihre Kenntnisse mit verhältnismäßig geringer Mühe auch auf andere Gebiete der Mathematik auszudehnen, wenn sie, wie es so oft vorkommt, das große Unrecht begehen, sich zu ausschließlich in Einzelheiten einzulassen, d. h. sich mit zu großer Ab-

schließung einem speziellen Teil der Wissenschaft zu widmen und alle übrigen Teile darüber zu vernachlässigen."

„Dieses Kultivieren einzelner Teile ist allmählich als notwendige Folge der ungeheuren Entwicklung eingetreten, welche in diesem Jahrhundert die verschiedenen Zweige der Wissenschaft erlebt haben. Aber auch in der Spezialisierung ist Maßhalten vonnöten, und ich habe immer geglaubt, daß es nicht schön ist, wenn man sieht, wie Mathematiker, man könnte sagen absichtlich, die eigene Ausbildung auf ein eng begrenztes Gebiet beschränken und sich für berechtigt halten, den benachbarten Gebieten keinen einzigen Blick zu schenken. Am wenigsten schön aber ist es, wenn junge Leute zu früh sich solchen Neigungen hingeben."

„Ich bin der Meinung, daß man mit allen Kräften eine so gefährliche Tendenz bekämpfen sollte. Eine geistige Diät, welche mit einer einzigen Speise beginnt und endet, kann niemandem zum Vorteil gereichen und nur große Übel zur Folge haben. Denn, wie Gladstone sehr richtig in einer Rede sagte, die er im Jahr 1879 an die Studenten der Universität Glasgow richtete, bei solcher Exklusivität beraubt man sich des Vorzugs des Seitenlichts, welches die Reiche der Wissenschaft gegenseitig aufeinanderwerfen, und man wird dazu neigen, den Wert und den Einfluß des eigenen beschränkten Verdienstes zu überschätzen. Wenn dieses Urteil nun für die Verhältnisse berechtigt ist, welche zwischen den verschiedenen Wissenschaften bestehen, um wieviel mehr wird es nicht für die verschiedenen Abteilungen und Unterabteilungen einer und derselben Wissenschaft gelten? Und muß es nicht noch viel mehr gerade bei den mathematischen Wissenschaften anerkannt werden, deren neueste Fortschritte immer deutlicher gezeigt haben, wie gebrechlich die Schranken sind, welche die verschiedenen Teile zu trennen schienen?"

„Die Anordnung des Stoffes ist bei jeder Theorie fast immer dieselbe; zuerst werden die Definitionen und Grundbegriffe der Theorie gegeben, alsdann die Theoreme und Formeln ohne Beweis aufgestellt, welche die Verbindung zwischen den durch die vorhergehenden Definitionen eingeführten Dingen oder Größen bilden, und schließlich ein kurzer Hinweis auf die Literatur über die betreffende Theorie gebracht."

„Da es nicht möglich war, alles zu bringen, habe ich mich sehr oft auf das Wichtigste beschränken müssen, und die

Schwierigkeiten, bei der Auswahl des Stoffes den richtigen Prinzipien zu folgen, sind nicht geringe und nicht wenige gewesen; ich glaube auch nicht, daß es mir immer gelungen ist, sie auf die beste Art zu überwinden; jedenfalls bitte ich den Leser, bei der Beurteilung aller Einzelheiten dieses bescheidenen Buches die größte ihm mögliche Nachsicht zu üben."

„Man glaube nicht, daß ich alle Literaturangaben gemacht hätte, die zu machen möglich waren; das wäre übertrieben und nutzlos gewesen; es war, wie mir scheint, nur geboten, die wichtigsten Arbeiten, die einen bestimmten Gegenstand betreffen, hervorzuheben, d. h. diejenigen, welche den größten Eindruck hinterlassen haben und als die Grundlagen der übrigen zu betrachten sind; denn wollte man sich von der Sucht, möglichst viel zu zitieren, beherrschen lassen, so würde dem Leser schließlich die natürlichste und einfachste Orientierung verloren gehen."

„Bei der allgemeinen Anordnung der verschiedenen Teile war ich manchmal gezwungen, um eine gewisse Symmetrie einhalten zu können, von der logischen Aufeinanderfolge abzuweichen und Theorien vorauszuschicken, zu deren Beweis (aber wohlverstanden nicht zum Verständnis der Resultate) Begriffe nötig sind, die erst später gebrachten Theorien angehören. Bei dem Charakter des Buches scheint mir dies kein Mißstand zu sein."

„Der Verfasser hofft, daß die mühsame Arbeit bei der Herstellung des Buches nützlich und nicht umsonst gewesen sei, und daß der nachsichtige Leser berücksichtigen werde, daß dieses Werk, in welchem die sämtlichen so verschiedenartigen Teile der reinen Mathematik behandelt werden, nicht das Resultat des Zusammenwirkens vieler Verschiedener, sondern die Arbeit eines einzigen ist."

ERNST PASCAL.

VORWORT ZUR ZWEITEN AUFLAGE.

In der oben wiedergegebenen Vorrede hatte der Verfasser als Zweck des Buches bezeichnet, „auf einem möglichst kleinen Raum die wichtigsten Theorien der neueren Mathematik zu vereinigen, von jeder Theorie nur so viel zu bringen, als nötig ist, damit der Leser sich in ihr orientieren könne und auf die Bücher zu verweisen, in welchen er Ausführlicheres finden kann". Diesem Zweck ist das Buch, dank dem bewundernswerten Fleiß und der großen Vielseitigkeit des Verfassers, unstreitig in hohem Maß gerecht geworden, und es hat die Beliebtheit, die es namentlich in den Kreisen der Studierenden genoß, wohl verdient. Trotzdem haben die Herausgeber der zweiten deutschen Auflage geglaubt, die Ziele des Buches erheblich höher und weiter stecken zu sollen. Sie sind dabei von der Erwägung ausgegangen, daß gerade in der Mathematik eine enzyklopädische Häufung von Einzelkenntnissen ohne inneren Zusammenhang weniger als irgendwo am Platze ist, daß vielmehr der angehende Mathematiker darauf Wert legen muß, *einen systematischen, auf wirklichem Verständnis beruhenden Überblick über das Gesamtgebiet der Wissenschaft zu gewinnen.* Ist hiermit die Hauptaufgabe gekennzeichnet, die das Repertorium in seiner jetzigen Gestalt zu erfüllen sucht, so möchte es daneben auch dem wissenschaftlich arbeitenden Mathematiker in knappen Umrissen ein Bild von dem heutigen Stand der Forschung geben und so dazu beitragen, die einzelnen Gebiete der Wissenschaft in lebendige Wechselbeziehung zu bringen. Diese veränderte Tendenz des Werkes machte eine völlige Umgestaltung der Form und des Inhalts notwendig, über welche in der Vorrede des zweiten Bandes eingehender berichtet ist. Daß sich dabei der Umfang des Buches sehr stark vermehrt hat, ließ sich im Interesse einer gleichmäßigen Behandlung der einzelnen Teile und bei Berücksichtigung der wichtigen Fortschritte des letzten Jahrzehnts nicht umgehen. Die Kapitel über Wahrscheinlichkeitsrechnung und mathematische Instrumente werden ausgeschieden. Sie sollen in einem in Aussicht genommenen

Repertorium der angewandten Mathematik Platz finden. Eine breitere Ausgestaltung haben die algebraisch-gruppentheoretischen Kapitel erhalten. Sie haben den Charakter von eingehenderen Referaten, die zur Orientierung auf einem Feld lebhafter mathematischer Tätigkeit vielfach willkommen sein werden. Besondere Sorgfalt wurde auf zweckmäßige Auswahl und Zuverlässigkeit der Literaturangaben verwendet. Sie sollen einen Überblick über den Entwicklungsgang der Wissenschaft gewähren und sind in solcher Vollständigkeit gegeben, daß man von ihnen aus leicht den Zugang zu der weiteren Literatur finden kann.

Stellt so das Buch in seiner jetzigen Gestalt ein ganz neues Werk vor, so haben die Herausgeber doch gern an die Spitze des Titels den Namen des verdienstvollen Verfassers der ersten Auflage gestellt. Er hat die Güte gehabt, zu der vorliegenden Auflage ein neues Kapitel über Differentialformen und totale Differentialgleichungen beizutragen.

Allen Herren, die durch ihre mühevolle Mitarbeit die Herausgabe des Werkes nach dem hier dargelegten Plan ermöglicht haben, sei auch an dieser Stelle herzlichster Dank ausgesprochen. Ganz besonders gilt dies Herrn A. Loewy in Freiburg, der von Anfang an sein lebhaftes Interesse an dem Buch gezeigt hat. Er hat fast sämtliche Druckbogen des ersten Bandes durchgesehen, und wir verdanken ihm viele wertvolle Anregungen und Zusätze.

Der Verlagsbuchhandlung sind wir für die sorgfältige Ausstattung des Buches und für das Entgegenkommen, das sie uns in gewohnter liebenswürdiger Weise bei der Drucklegung gezeigt hat, zu großem Dank verpflichtet.

Strassburg i. E., Ostern 1910.

PAUL EPSTEIN.

INHALTSVERZEICHNIS.

Kapitel I.

Arithmetik, Mengenlehre, Grundbegriffe der Funktionenlehre.

Von *Hans Hahn* in Czernowitz.

		Seite
§ 1.	Größensysteme. Natürliche und rationale Zahlen	1
§ 2.	Irrationale Zahlen	6
§ 3.	Die gemeinen komplexen Zahlen	11
§ 4.	Die höheren komplexen Zahlen	14
§ 5.	Grundbegriffe der Mengenlehre. Die Mächtigkeiten der Mengen	17
§ 6.	Die geordneten Mengen. Die Ordnungstypen	20
§ 7.	Die wohlgeordneten Mengen	23
§ 8.	Die Punktmengen	26
§ 9.	Der Funktionsbegriff. Der Begriff des Grenzwertes	30
§ 10.	Obere und untere Grenze. Stetigkeit und Unstetigkeit	34
§ 11.	Potenzen. Logarithmen. Wichtige Grenzwerte	39

Kapitel II.

Kombinatorik, Determinanten und Matrices.

Von *Alfred Loewy* in Freiburg i. Br.

§ 1.	Die Lehre von den Kombinationen. Die Binomialkoeffizienten. Die Polygonal- und Pyramidalzahlen. Die figurierten Zahlen. Die Polynomialkoeffizienten	43
§ 2.	Historisches und Allgemeines über Determinanten	52
§ 3.	Symmetrische und schiefe Determinanten. Jacobische Symbole. Hermitesche Determinanten	63
§ 4.	Spezielle Determinanten	66
§ 5.	Systeme linearer Gleichungen	72
§ 6.	Das Rechnen mit Matrices. Zusammenhang zwischen Matrices, linearen Substitutionen und bilinearen Formen	79
§ 7.	Bilineare Formen und höhere komplexe Zahlen	93
§ 8.	Die Theorie der Elementarteiler und die bilinearen Formen	102
§ 9.	Die quadratischen und Hermiteschen Formen. Die alternierenden bilinearen Formen	118
§ 10.	Orthogonale Substitutionen und automorphe Transformationen quadratischer, Hermitescher und bilinearer Formen	130
§ 11.	Abgeleitete Matrices. Frankesche und Sylvestersche Sätze. Laplacescher und Jacobischer Satz. Potenztransformation. Produkttransformation. Lineare homogene Substitutionen, die eine Determinante in sich überführen	138
§ 12.	Differentiation einer Determinante und Matrix	153
§ 13.	Die Wronskische Determinante	154

Seite
§ 14. Die Jacobische oder Funktionaldeterminante 156
§ 15. Die Hessesche Determinante 160
§ 16. Unendliche Determinanten. Systeme von unendlich vielen linearen Gleichungen. Bilineare Formen mit unendlich vielen Variablen 161
§ 17. Kubische Determinanten und solche höheren Ranges . . 167

Kapitel III.

Algebraische Gruppentheorie.

Von *Alfred Loewy* in Freiburg i. Br.

§ 1. Historisches. Definition der Gruppe. Beispiele. Das Feld oder der Körper 168
§ 2. Allgemeines über abstrakte Gruppen 180
§ 3. Abstrakte endliche Gruppen 194
§ 4. Auflösbare, einfache und zusammengesetzte Gruppen . . 199
§ 5. Abelsche Gruppen. Hamiltonsche Gruppen. Die Quaternionengruppe 202
§ 6. Permutationen. Symmetrische und alternierende Gruppe 204
§ 7. Transitive, intransitive, primitive und imprimitive Permutationsgruppen. Reguläre Gruppen. Gruppen vom Primzahlgrad. Metazyklische Gruppe. Modulargruppe . . . 209
§ 8. Permutationsgruppen höchster Ordnung bei gegebenem Grad. Die zu den Permutationsgruppen zugehörigen Funktionen. Die symmetrischen und alternierenden Funktionen 216
§ 9. Lineare homogene Substitutionsgruppen. Reduzibilität. Endliche und unendliche Gruppen. Invarianten endlicher Gruppen 222
§ 10. Darstellung einer endlichen abstrakten Gruppe als Permutationsgruppe und als Gruppe linearer homogener Substitutionen 228
§ 11. Kollineationsgruppen. Darstellbarkeit einer endlichen abstrakten Gruppe als Kollineationsgruppe. Die verschiedenen Typen endlicher Kollineationsgruppen in einer, zwei und drei Variablen 233
§ 12. Lineare Gruppen mit Koeffizienten aus einem Körper. Kongruenzgruppen 245

Kapitel IV.

Algebraische Gleichungen.

Von *Alfred Loewy* in Freiburg i. Br.

§ 1. Historisches. Literatur 250
§ 2. Rationale Funktionen. Teilbarkeitsgesetze 257
§ 3. Fundamentalsatz der Algebra. Einfache und mehrfache Wurzeln 260
§ 4. Partialbruchzerlegung einer gebrochenen rationalen Funktion 262
§ 5. Symmetrische Funktionen der Gleichungswurzeln 265

Seite

§ 6. Resultante und Diskriminante 268
§ 7. Transformation einer Gleichung. Tschirnhausentransformation. Mit der Tschirnhausentransformation zusammenhängende Normalformen. Funktionen mehrerer Wurzeln. Gleichung der quadrierten Wurzeldifferenzen 276
§ 8. Die kubische und die biquadratische Gleichung. Reziproke Gleichungen 281
§ 9. Körper. Reduzibilität und Irreduzibilität. Algebraischer Körper. Primitive und imprimitive Größen. Normalgleichung 290
§ 10. Die Galoissche Gruppe einer Gleichung. Natürliche und akzessorische Irrationalitäten. Rationale Resolventen. Reduktion der Gruppe und Rückführung einer Gleichung auf eine Kette von Hilfsgleichungen. Allgemeine Gleichungen. Affektlose Gleichungen 298
§ 11. Abelsche Gleichungen 305
§ 12. Die Einheitswurzeln. Die Kreisteilungsgleichung. Gaußsche Summen. Reine Gleichung 311
§ 13. Algebraisch auflösbare Gleichungen 320
§ 14. Gleichungen fünften und höheren Grades. Einige funktionentheoretische und geometrische, nicht algebraisch auflösbare Gleichungen höheren Grades 326
§ 15. Bestimmung der Anzahl der reellen und komplexen Wurzeln einer Gleichung 337
§ 16. Approximation der Wurzeln 349

Kapitel V.

Invariantentheorie.

Von *H. E. Timerding* in Braunschweig.

§ 1. Binärformen zweiter bis vierter Ordnung 358
§ 2. Binärformen beliebig hoher Ordnung 366
§ 3. Die Cayley-Aronholdsche Symbolik 376
§ 4. Kanonische und typische Darstellung. Erweiterungen des Invariantenbegriffes 387
§ 5. Formenmoduln 393
§ 6. Invariante Bildungen der Formen mit mehr als zwei Veränderlichen 403
§ 7. Besondere ternäre Systeme 411

Kapitel VI.

Reihen, Produkte, Kettenbrüche.

Von *Paul Epstein* in Straßburg i. E.

§ 1. Endliche Reihen 421
§ 2. Unendliche Reihen 424
§ 3. Reihen von Funktionen 431
§ 4. Mehrfache Reihen 439
§ 5. Unendliche Produkte 442
§ 6. Allgemeine Kettenbrüche 444
§ 7. Regelmäßige Kettenbrüche 449

Kapitel VII.

Differentialrechnung.

Von *Paul Epstein* in Straßburg i. E.

		Seite
Einleitung		454
§ 1.	Infinitesimale Größen	456
§ 2.	Begriff und Existenz des Differentialquotienten	457
§ 3.	Grundregeln für die Differentiation entwickelter Funktionen einer Veränderlichen	460
§ 4.	Differentialquotienten höherer Ordnung von Funktionen einer Veränderlichen	462
§ 5.	Differentiation von Funktionen mehrerer Veränderlichen, von zusammengesetzten und unentwickelten Funktionen einer Veränderlichen	465
§ 6.	Rollesches Theorem. Mittelwertsätze. Taylorscher Lehrsatz	469
§ 7.	Grenzwerte von Ausdrücken in unbestimmter Form	473
§ 8.	Wachsende und abnehmende Funktionen. Maxima und Minima	474

Kapitel VIII.

Integralrechnung.

Von *Paul Epstein* in Straßburg i. E.

§ 1.	Das unbestimmte Integral. Definition und allgemeine Sätze	478
§ 2.	Integration wichtiger Klassen von Funktionen	480
§ 3.	Das bestimmte Integral. Definitionen	486
§ 4.	Haupteigenschaften der bestimmten Integrale	494
§ 5.	Zusammenstellung einiger bestimmter Integrale	497
§ 6.	Kurvenintegrale	502
§ 7.	Mehrfache Integrale	506

Kapitel IX.

Differenzenrechnung.

Von *H. E. Timerding* in Braunschweig.

§ 1.	Differenzen	511
§ 2.	Interpolation	514
§ 3.	Summen	517
§ 4.	Zusammenhang zwischen Summen und Integralen. Mechanische Quadratur	521

Verzeichnis der hauptsächlichsten Abkürzungen,

die bei den Zitaten angewandt sind.

Acta Math.	Acta Mathematica (Stockholm)
Amer. J.	American Journal of Mathematics (Baltimore)
Ann. di Mat.	Annali di Matematica pura ed applicata (Milano)
Ann. of Math.	Annals of Mathematics (Cambridge, Mass.)
Ann. éc. norm.	Annales scientifiques de l'école normale supérieure (Paris)
Arch. Math. Phys.	Archiv der Mathematik und Physik (Leipzig)
Berl. Abh.	Abhandlungen der kgl. preußischen Akademie der Wissenschaften in Berlin
Berl. Monatsber.	Monatsberichte der kgl. preußischen Akademie der Wissenschaften in Berlin
Berl. Sitzungsber.	Sitzungsberichte der kgl. preußischen Akademie der Wissenschaften in Berlin
Bibl. Math.	Bibliotheca Mathematica (Leipzig)
Bull. Am. M. S.	Bulletin of the American Mathematical Society (New York)
Bull. Sc. M.	Bulletin des sciences mathématiques
Bull. Soc. M.	Bulletin de la Société Mathématique de France (Paris)
C. R.	Comptes rendus hebdomadaires des séances de l'Académie des Sciences (Paris)
Dublin Proc.	Proceedings of the Royal Irish Academy
Dublin Trans.	Transactions of the Royal Irish Academy
Edinb. M. S. Proc.	Proceedings of the Edinburgh Mathematical Society
Edinb. R. S. Proc.	Proceedings of the Royal Society of Edinburgh
Edinb. Trans.	Transactions of the Royal Society of Edinburgh
Enzykl.	Enzyklopädie der mathematischen Wissenschaften
Fortschr. d. Math.	Jahrbuch über die Fortschritte der Mathematik
Giorn. di Mat.	Giornale di Matematiche di Battaglini (Napoli)
Gött. Abh.	Abhandlungen der kgl. Gesellschaft der Wissenschaften zu Göttingen
Gött. Nachr.	Nachrichten der kgl. Gesellschaft der Wissenschaften zu Göttingen
J. éc. polyt.	Journal de l'école polytechnique (Paris)
J. de Math.	Journal de mathématiques pures et appliquées (Paris)
J. f. Math.	Journal für die reine und angewandte Mathematik, gegründet von A. L. Crelle (Berlin)
Ist. Lomb. Rend.	Reale Istituto Lombardo di Scienze e Lettere. Rendiconti (Milano)
Leipz. Abh.	Abhandlungen der kgl. sächsischen Gesellschaft der Wissenschaften in Leipzig
Leipz. Ber.	Berichte über die Verhandlungen der kgl. sächsischen Gesellschaft der Wissenschaften in Leipzig

Lond. M. S. Proc.	Proceedings of the London Mathematical Society
Lond. R. S. Proc.	Proceedings of the Royal Society of London
Math. Ann.	Mathematische Annalen, begründet 1868 durch A. Clebsch (Leipzig)
Math.-Ver.	Jahresberichte der deutschen Mathematiker-Vereinigung (Leipzig)
Mess. of Math.	The Messenger of Mathematics (London)
Monatsh. f. Math.	Monatshefte für Mathematik und Physik (Wien)
Münch. Abh.	Abhandlungen der kgl. bayerischen Akademie der Wissenschaften zu München
Münch. Ber.	Sitzungsberichte der math.-phys. Klasse der kgl. bayerischen Akademie der Wissenschaften zu München
Nouv. Ann.	Nouvelles Annales de Mathématiques (Paris)
(*Nov.*) *Comm. Petrop.*	(Novi) Commentarii Academiae Scientiarum Imperialis Petropolitanae
Pal. Circ. mat.	Rendiconti del Circolo matematico di Palermo
Phil. Mag.	The London, Edinburgh and Dublin philosophical magazine and journal of Science
Phil. Trans.	Philosophical Transactions of the Royal Society of London
Prag. Ber.	Sitzungsberichte der kgl. böhmischen Gesellschaft der Wissenschaften
Quart. J.	The Quarterly Journal of pure and applied mathematics (London)
Rev. de Math.	Revue de Mathématiques (Rivista di Matematica) publiée par G. Peano (Turin)
Rev. de Math. spéc.	Revue de Mathématiques spéciales (Paris)
Rom. Acc. L. Mem.	Memorie della R. Accademia dei Lincei
Rom. Acc. L. Rend.	Atti della R. Accademia dei Lincei. Rendiconti
Torino Atti	Atti della R. Accademia delle scienze di Torino
Torino Mem.	Memorie della R. Accademia delle scienze di Torino
Trans. Am. M. S.	Transactions of the American Mathematical Society (Lancaster und New York)
Zschr. Math. Phys. (*Zschr. f. Math.*)	Zeitschrift für Mathematik und Physik, begründet 1856 durch O. Schlömilch (Leipzig)
Zschr. f. math. Unt.	Zeitschrift für mathematischen und naturwissenschaftlichen Unterricht (Leipzig)

Die eingeklammerten Ortsnamen geben in einzelnen Fällen den gegenwärtigen Erscheinungsort an. Die Akademieschriften beziehen sich überall, wo mehrere Abteilungen vorhanden, auf die mathematisch-physikalische Klasse.

Eine eingeklammerte Ziffer bezeichnet in den Zitaten die Serie, eine fette Zahl den Band, eine gewöhnliche Zahl die Seite der periodischen Veröffentlichung.

Kapitel I.

Arithmetik, Mengenlehre, Grundbegriffe der Funktionenlehre.

Von *Hans Hahn* in Wien.

§ 1. Größensysteme. Natürliche und rationale Zahlen.

Unter einem *Größensystem* versteht man, nach der Definition von H. Grassmann (*Lehrbuch der Arithmetik*, Berlin 1861, *Werke* **2**, 1. Teil, 298), ein System von Dingen, von denen zwei beliebige (a und b) auf Grund irgendeiner Regel entweder als einander gleich ($a = b$ oder $b = a$) oder als ungleich ($a \neq b$ oder $b \neq a$) erklärt sind, vorausgesetzt, daß die Erklärung der Gleichheit den folgenden Bedingungen genügt: 1. Die Relationen $a = b$ und $a \neq b$ schließen sich gegenseitig aus; 2. jede Größe ist sich selbst gleich ($a = a$); 3. wenn zwei Größen derselben dritten Größe gleich sind, so sind sie untereinander gleich (neben $a = b$, $a = c$ ist auch $b = c$).

Ein Größensystem heißt geordnet, wenn von irgend zwei ungleichen Dingen desselben (a und b) auf Grund irgendeiner Regel das eine (etwa a) als das größere, das andere als das kleinere erklärt ist ($a > b$ oder $b < a$), vorausgesetzt, daß die Erklärung den folgenden Bedingungen genügt: 1. Die Relationen $a > b$ und $a < b$ schließen sich gegenseitig aus; 2. aus $a > b$ und $a = a'$, $b = b'$ folgt $a' > b'$; 3. aus $a > b$ und $b > c$ folgt $a > c$.

Unter einer *Verknüpfung* (Operation) in unserem Größensysteme verstehen wir irgendeine Regel, die zwei Dingen (a und b) des Systemes ein drittes (sowie jedes diesem dritten Dinge gleiche Ding des Systemes) zuordnet, sei es, daß a und b noch gewissen Beschränkungen unterliegen, sei es, daß sie in unserem Größensysteme beliebig gewählt werden können (im letzteren Falle sagt man, die Verknüpfung sei im betrachteten Größensysteme allgemein ausführbar). Das Resultat

der Verknüpfung von a mit b werde bezeichnet durch $a \circ b$; soll dieses Verknüpfungsresultat selbst wieder mit einer Größe verknüpft werden, so setzt man es in Klammern: $(a \circ b)$. Von den Verknüpfungen ist stets vorauszusetzen, daß aus $a = a'$ und $b = b'$ folgt: $a \circ b = a' \circ b'$.

Eine Verknüpfung heißt *assoziativ,* wenn

$$(1) \qquad (a \circ b) \circ c = a \circ (b \circ c),$$

sie heißt *kommutativ,* wenn

$$(2) \qquad a \circ b = b \circ a.$$

Ist die betrachtete Verknüpfung assoziativ, so bleibt der Ausdruck $\{[(a_1 \circ a_2) \circ a_3] \cdots \circ a_{n-1}\} \circ a_n$ *ungeändert, wenn man in ihm die einzelnen Größen statt in der angegebenen Weise in beliebiger anderer Weise durch Klammern zusammenfaßt.* Man kann diesen Ausdruck also auch mit Hinweglassung aller Klammern in der Form schreiben: $a_1 \circ a_2 \circ a_3 \cdots \circ a_n$.

Ist die betrachtete Verknüpfung assoziativ und kommutativ, so ändert der Ausdruck $a_1 \circ a_2 \circ a_3 \cdots \circ a_n$ *seinen Wert nicht, wenn man in ihm die Reihenfolge der einzelnen Glieder beliebig abändert.*

Sei in unserem Größensysteme noch eine zweite Verknüpfung definiert, die mit $\odot$ bezeichnet werde; sie heißt *distributiv* zur ersten Verknüpfung, wenn

$$(3) \quad a \odot (b \circ c) = (a \odot b) \circ (a \odot c) \text{ und } (b \circ c) \odot a = (b \odot a) \circ (c \odot a).$$

Ist die Verknüpfung $\odot$ kommutativ, so reduzieren sich diese beiden Gleichungen auf eine.

Wir setzen von nun an voraus, die durch $\circ$ bezeichnete Operation sei allgemein ausführbar, assoziativ und kommutativ. Sind a und b zwei Größen des Systemes, so wird die Aufsuchung einer der Gleichung $b \circ x = a$ genügenden Größe als die zu unserer Verknüpfung *inverse Verknüpfung* bezeichnet (oder auch als *lytische Operation;* im Gegensatz hierzu heißt die zuerst definierte Verknüpfung auch *thetische Operation*). Wir drücken die inverse Verknüpfung aus durch $a \cup b$. Gibt es eine der genannten Gleichung genügende Größe x, so heißt die lytische Operation $a \cup b$ ausführbar, sind alle dieser Gleichung genügenden Größen einander gleich, so heißt sie *eindeutig ausführbar* (ist die thetische Operation nicht kommutativ, so hat man zwei lytische Operationen zu unterscheiden, eine *vordere*

und eine *hintere*, je nachdem es sich um Auflösung der Gleichung $x \circ b = a$ oder $b \circ x = a$ handelt).

Ist unser Größensystem ein geordnetes und genügt die thetische Operation der Bedingung:

(4) *Aus* $d > d'$ *folgt* $c \circ d > c \circ d'$,

so ist die lytische Operation $a \cup b$ *immer, wenn sie ausführbar ist, auch eindeutig.*

Ist in unserem Größensysteme die zur betrachteten Operation inverse zwar nicht allgemein ausführbar, aber immer, wenn sie ausführbar ist, eindeutig, so läßt sich das Größensystem durch Hinzufügung neuer Größen so erweitern und auch die betrachtete Operation auf das neue Größensystem so erweitern, daß sie assoziativ und kommutativ bleibt, die zu ihr inverse Operation aber allgemein eindeutig ausführbar wird. War das ursprüngliche System geordnet und genügte unsere Operation der Bedingung (4), so läßt sich diese Ordnung auch auf das neue System erweitern, und zwar so, daß auch die erweiterte Operation der Bedingung (4) genügt.

Diese Erweiterung geschieht in folgender Weise. Man fügt den Größen des ursprünglichen Systemes als neue Dinge die Paare (a, b) solcher Größen hinzu und definiert die Gleichheit durch die Festsetzungen: 1. Das Paar (a, b) sei gleich dem Paare (a', b'), wenn $a \circ b' = a' \circ b$; 2. das Paar (a, b) sei gleich der Größe c des ursprünglichen Systemes, wenn $a \cup b$ im ursprünglichen Systeme ausführbar ist und c ergibt. Dann ist die beliebige Größe c des ursprünglichen Systemes gleich dem Paare $(a \circ c, a)$, so daß wir alle folgenden Beziehungen nur für Größenpaare festzusetzen brauchen.

Die Erweiterung unserer Operation wird definiert durch:

$$(a, b) \circ (a', b') = (a \circ a', b \circ b').$$

Dann wird:

$$(a, b) \cup (a', b') = (a \circ b', a' \circ b);$$

insbesondere wird:

$$a \cup a' = (a, a').$$

Die Erweiterung der Ordnung geschieht durch die Festsetzung:

$$(a, b) > (a', b') \quad \text{wenn} \quad a \circ b' > b \circ a'.$$

Im erweiterten Systeme gibt es eine und nur eine Größe, die mit einer beliebigen Größe dieses Systemes verknüpft dieselbe

ungeändert läßt; sie heißt indifferente Größe (oder Modul) und ist gleich dem Paare (a, a), *oder was dasselbe ist, dem Ausdruck* $a \cup a$, *wo* a *eine beliebige Größe des ursprünglichen Systemes ist.* Die indifferente Größe kann auch schon im ursprünglichen Systeme enthalten sein.

Dasjenige Größensystem, auf welchem sich die Arithmetik aufbaut, ist das System der *natürlichen* (ganzen positiven) *Zahlen.* Mit einer Analyse des Begriffes der natürlichen Zahl haben sich viele Autoren beschäftigt. Wir nennen die wichtigsten dieser Schriften: E. Schröder, *Lehrbuch der Arithmetik und Algebra*, Leipzig 1873, H. v. Helmholtz, *Zählen und Messen, Wissensch. Abh.* **3**, 356, L. Kronecker, *Über den Zahlbegriff, Journ. f. Math.* **101**, 337 (diese beiden Abhandlungen sind zuerst erschienen in: *Philosophische Aufsätze, E. Zeller zu seinem 50jährigen Doktorjubiläum gewidmet,* Leipzig 1887). R. Dedekind, *Was sind und was sollen die Zahlen?* Braunschweig 1888, G. Peano, *Arithmetices principia nova methodo exposita,* Turin 1889, *Sul concetto di numero, Rev. de math.* **1**, 87, *Formulaire de math.* **2**, § 2 (Turin 1898), **3**, 39 (1901), **4**, 53 (1903); siehe auch Genocchi-Peano, *Differentialrechnung und Grundzüge der Integralrechnung* (deutsch von Bohlmann u. Schepp), 336, Leipzig 1899, G. Frege, *Grundgesetze der Arithmetik,* Jena 1893 u. 1903, D. Hilbert, *Über die Grundlagen der Logik und der Arithmetik, Verh. d. 3. intern. Math.-Kongr. 1904,* 174, B. Russell, *The principles of mathematics* **1**, Cambridge 1903, L. Couturat, *Les principes des mathématiques,* Paris 1905 und *Rev. d. metaph. et d. mor.* **12**, 19, 211 (1904), H. Weber, *Elementare Mengenlehre, Math.-Ver.* **15**, 173 (1906) (abgedr. in H. Weber und J. Wellstein, *Enzykl. d. Elem.-Math.*, Leipzig (1907), **3**, 645; vgl. auch P. Stäckel, *Math.-Ver.* **16**, 425).

Im Gebiete der natürlichen Zahlen gilt der Satz von der *vollständigen Induktion: Es sei eine Aussage, in der eine unbestimmte natürliche Zahl* n *vorkommt, als richtig erkannt für* $n = 1$, *und aus der Annahme ihrer Richtigkeit für eine bestimmte Zahl* n *folge ihre Richtigkeit für die nächstgrößere Zahl* $n + 1$; *dann ist die Aussage richtig für alle natürlichen Zahlen.*

Die *Addition* und die *Multiplikation* der natürlichen Zahlen sind unbeschränkt ausführbar, assoziativ und kommutativ, und genügen der Bedingung (4). Die Multiplikation ist in Verbindung mit der Addition distributiv. Die umgekehrten Operationen. *Subtraktion* und *Division*, sind im Gebiete der natür-

lichen Zahlen nicht unbeschränkt ausführbar, soweit sie aber ausführbar sind, eindeutig.

Wir machen zunächst die Division allgemein ausführbar, indem wir, nach der obigen Theorie, das System der natürlichen Zahlen erweitern durch Hinzufügung der Paare natürlicher Zahlen, die nichts anderes sind als die (positiven oder absoluten) Brüche (oder positiven rationalen Zahlen). Für ihre Multiplikation und Division ergibt diese Theorie die Regeln:

$$\frac{a}{b} \cdot \frac{a'}{b'} = \frac{aa'}{bb'}; \quad \frac{a}{b} : \frac{a'}{b'} = \frac{ab'}{a'b}.$$

Es läßt sich nun auch die Addition auf unser neues Zahlensystem so erweitern, daß sie assoziativ und kommutativ bleibt und in Verbindung mit ihr die Multiplikation distributiv ist. Es geschieht durch die Formel:

$$\frac{a}{b} + \frac{a'}{b'} = \frac{ab' + a'b}{bb'},$$

und zwar läßt sich zeigen, daß dies die einzig mögliche Erweiterung der Addition ist, der die genannten Eigenschaften zukommen. Sie genügt auch der Bedingung (4); die Subtraktion ist nicht allgemein ausführbar, aber wenn ausführbar, eindeutig.

Um die Subtraktion allgemein ausführbar zu machen, fügen wir wieder zu den bisherigen Zahlen die aus ihnen gebildeten Paare hinzu, wodurch man zum System der allgemeinen (positiven und negativen) rationalen Zahlen gelangt. Unter ihnen befindet sich die gegenüber der Addition indifferente Größe, die Null. Addition und Subtraktion wird gleichfalls durch unsere obige Theorie geliefert. Jede der neu eingeführten Zahlen (außer der Null) läßt sich auf die Form bringen: $0 - a$, wo a positiv rational, und wird dann mit $(-a)$ bezeichnet. Nun läßt sich wieder die Multiplikation auf das Gesamtgebiet der rationalen Zahlen erweitern, so daß sie assoziativ und kommutativ bleibt und in Verbindung mit der Addition distributiv wird. Es geschieht dies durch die Regel: *Ein Produkt, von dem wenigstens ein Faktor Null ist, ist Null,* und durch die Regeln:

$$(-a) \cdot b = -(ab); \quad (-a) \cdot (-b) = ab.$$

Wieder läßt sich zeigen, dass diese Erweiterung der Multiplikation die einzige ist, der die genannten Eigenschaften zukommen. Doch genügt sie nicht der Bedingung (4). Die Division ist eindeutig ausführbar, außer wenn der Divisor Null ist. In

diesem Falle ist sie unausführbar (wenn der Dividend nicht Null ist) oder unendlich vieldeutig (wenn auch der Dividend Null ist).

Im Gebiete der rationalen Zahlen sind also Addition, Multiplikation, Subtraktion und Division unbeschränkt und eindeutig ausführbar, mit Ausnahme der Division durch Null.

Eine nähere Ausführung der im obigen skizzierten Theorie findet man bei O. Stolz und J. A. Gmeiner, *Theoretische Arithmetik*, Leipzig 1900 (*Teubners Sammlung* 4). Ferner L. Couturat, *De l'infini mathématique* 1 ff. (Paris 1896) und die oben genannten Schriften von G. Peano.

§ 2. Irrationale Zahlen.

Es sei eine Einteilung der rationalen Zahlen in zwei Klassen A und B gegeben, derart daß:

1. jede rationale Zahl entweder zu A oder zu B gehört,
2. sowohl A als B rationale Zahlen enthalten, und
3. jede in A stehende Zahl kleiner ist als jede in B stehende Zahl.

Eine solche Einteilung der rationalen Zahlen in zwei Klassen heißt ein *Schnitt im Gebiete der rationalen Zahlen.* Wir bezeichnen ihn kurz als den Schnitt (A, B). Sei c eine rationale Zahl; gibt man eine rationale Zahl a in die Klasse A, wenn $a < c$ (oder $a \leqq c$) und eine rationale Zahl b in die Klasse B, wenn $b \geqq c$ (oder $b > c$), so erhält man einen solchen Schnitt; er heißt *erzeugt durch die rationale Zahl c.* Es gibt aber auch Schnitte im Gebiete der rationalen Zahlen, welche durch keine rationale Zahl erzeugt werden. Von einem solchen Schnitt sagt man, *er definiert eine irrationale Zahl* γ oder, *er wird erzeugt durch die irrationale Zahl* γ. Die Gesamtheit aller rationalen und irrationalen Zahlen bildet das Gebiet der reellen Zahlen. Ist γ definiert durch den Schnitt (A, B), so heißt γ größer als jede in A stehende rationale Zahl und kleiner als jede in B stehende rationale Zahl. Die Zahl γ heißt positiv oder negativ, je nachdem sie größer oder kleiner als Null ist.

Zwei durch die Schnitte (A, B) und (A', B') definierte irrationale Zahlen γ und γ' heißen *gleich,* wenn die Klasse rationaler Zahlen A identisch ist mit A', und ebenso B identisch mit B'. Die Zahl γ heißt *größer als* γ', wenn eine in A stehende rationale Zahl in B' vorkommt; γ heißt *kleiner als* γ', wenn eine in B stehende rationale Zahl in A' vorkommt.

Seien γ und γ' die die beiden Schnitte (A, B) und (A', B') erzeugenden reellen Zahlen und es mögen mit a, b, a', b' die bezw. in A, B, A', B' stehenden rationalen Zahlen bezeichnet werden. Fassen wir alle Zahlen $a + a'$ in eine, alle Zahlen $b + b'$ in eine andere Klasse zusammen, so bilden diese beiden Klassen einen Schnitt im Gebiete der rationalen Zahlen. Die ihn erzeugende reelle Zahl wird als die *Summe* $\gamma + \gamma'$ der beiden reellen Zahlen γ und γ' bezeichnet. Sind γ und γ' rational, so ist ihre so definierte Summe identisch mit ihrer Summe im gewöhnlichen Sinne.

Die eben definierte Addition der reellen Zahlen ist *assoziativ* und *kommutativ*. Sie läßt ferner eine eindeutige Umkehrung, die *Subtraktion*, zu, d. h. sind α und β reelle Zahlen, so gibt es eine und nur eine reelle Zahl ξ, so daß $\beta + \xi = \alpha$ ist; sie wird bezeichnet mit $\alpha - \beta$.

Die Zahl $0 - \alpha$ heißt die zu α entgegengesetzte reelle Zahl und wird mit $-\alpha$ bezeichnet; sie ist negativ oder positiv, je nachdem α positiv oder negativ ist. Unter dem *absoluten Betrage* (oder *Modul*) *von* α, in Zeichen $|\alpha|$, versteht man die Zahl α selbst, oder die entgegengesetzte, je nachdem α positiv oder negativ ist. Für $\alpha \neq 0$ ist daher $|\alpha| > 0$.

Um die *Multiplikation* zweier reeller Zahlen γ und γ' zu definieren, setzen wir zunächst γ und γ' als positiv voraus. Faßt man dann alle Zahlen $b \cdot b'$ in eine Klasse, alle übrigen rationalen Zahlen in eine andere Klasse zusammen, so bilden diese beiden Klassen einen Schnitt; die ihn erzeugende reelle Zahl, die sicher positiv ist, wird als das *Produkt* $\gamma \cdot \gamma'$ der beiden Zahlen γ und γ' bezeichnet. Um die Multiplikation auch für nicht positive Zahlen zu definieren, setzt man fest: Ist wenigstens einer der beiden Faktoren Null, so ist auch das Produkt Null. Sind beide Faktoren von Null verschieden, so ist der absolute Betrag des Produktes gleich dem Produkt der absoluten Beträge, das Vorzeichen aber das positive oder das negative, je nachdem die beiden Faktoren gleiches oder ungleiches Zeichen haben.

Die so definierte Multiplikation der reellen Zahlen, die, wenn γ und γ' beide rational sind, sich auf die bekannte Multiplikation der rationalen Zahlen reduziert, ist *assoziativ*, *kommutativ* und in Verbindung mit der Addition *distributiv*. Ist $\beta \neq 0$, so gibt es eine und nur eine reelle Zahl ξ, so daß $\beta \cdot \xi = \alpha$ wird. Sie wird bezeichnet mit $\alpha : \beta$ oder $\frac{\alpha}{\beta}$.

Die *Division* ist also stets eindeutig ausführbar, außer durch Null.

Sind irgend zwei voneinander verschiedene reelle Zahlen α und β gegeben, so gibt es sowohl unendlich viele rationale, als auch unendlich viele irrationale Zahlen, die der Größe nach zwischen α und β liegen.

Man sagt von irgendwelchen reellen Zahlen $a_1, a_2, \ldots, a_n, \ldots$, deren Index n alle natürlichen Zahlen durchläuft: sie bilden eine *Folge reeller Zahlen.* Wir bezeichnen eine solche Folge kurz mit (a_n).

Die Folge (a_n) heißt *konvergent,* wenn zu jeder positiven Zahl ε ein Wert N des Index n gehört, derart, daß für alle Indices n' und n'', die größer als N sind, die Ungleichung besteht:

$$|a_{n'} - a_{n''}| < \varepsilon.$$

Eine Folge, die nicht konvergent ist, heißt *divergent.*

Die Zahl a heißt *Grenze* oder *Limes der Folge* (a_n), in Zeichen

$$a = \lim_{n=\infty} a_n,$$

wenn zu jedem positiven ε ein N gehört, derart, daß für jeden Index n, der größer als N ist, die Ungleichung besteht:

$$|a - a_n| < \varepsilon.$$

Jede Folge, die eine Zahl zur Grenze hat, ist konvergent. Umgekehrt: *Jede konvergente Folge hat eine und nur eine Zahl zur Grenze.*

Weiteres über den Grenzbegriff findet sich in § 9.

Sei e eine von 1 verschiedene positive ganze Zahl. Dann heißt der Ausdruck:

$$(1) \qquad d_n = \frac{c_1}{e} + \frac{c_2}{e^2} + \cdots + \frac{c_n}{e^n},$$

in dem $c_1, c_2, \ldots, c_n$ ganze Zahlen bedeuten, die zwischen 0 und $e - 1$ liegen (diese Grenzen eingeschlossen), *ein systematischer Bruch von der Grundzahl e.*

Sei nun $c_1, c_2, \ldots, c_n, \ldots$ eine Folge ganzer Zahlen, die zwischen 0 und $e - 1$ liegen (die Grenzen eingeschlossen). Wird dann die Zahl d_n durch Gleichung (1) definiert, so ist die Folge (d_n) stets konvergent. Bezeichnet d die Grenze von (d_n), so nennt man d die Summe des *unendlichen systematischen Bruches:*

$$\frac{c_1}{e} + \frac{c_2}{e^2} + \cdots + \frac{c_n}{e^n} + \cdots = \sum_{n=1}^{\infty} \frac{c_n}{e^n}.$$

Jede beliebige reelle Zahl zwischen 0 und 1 läßt sich in einen (endlichen oder unendlichen) *systematischen Bruch von gegebener Grundzahl e entwickeln* (dabei bedeutet e eine ganze Zahl, größer als 1). Genauer: sei d eine reelle Zahl zwischen 0 und 1. Ist dann d rational, so kann es auf eine und nur eine Weise in die Form gebracht werden $d = \frac{p}{q}$, wo p und q teilerfremde ganze positive Zahlen sind. Enthält dann q nur solche Primfaktoren, die auch in e aufgehen, so gibt es einen und nur einen *endlichen* systematischen Bruch von der Grundzahl e, dessen Summe gleich d ist, etwa:

$$\frac{c_1}{e} + \frac{c_2}{e^2} + \cdots + \frac{c_n}{e^n},$$

und außerdem noch einen und nur einen *unendlichen* solchen systematischen Bruch von der Summe d, nämlich:

$$d = \frac{c_1}{e} + \frac{c_2}{e^2} + \cdots + \frac{c_{n-1}}{e^{n-1}} + \frac{c_n - 1}{e^n} + \frac{e-1}{e^{n+1}} + \frac{e-1}{e^{n+2}} + \cdots.$$

Ist hingegen d irrational oder eine rationale Zahl, die nicht die genannte spezielle Form hat, so gibt es einen und nur einen *unendlichen* systematischen Bruch von der Grundzahl e, dessen Summe d ist.

Damit die Summe eines unendlichen systematischen Bruches eine rationale Zahl sei, ist notwendig und hinreichend, daß die Folge der Koeffizienten c_n von einem bestimmten Index an periodisch sei (d. h. es gibt ein N und ein k, so daß für $n > N$ immer $c_{n+k} = c_n$ ist).

Näheres über systematische Brüche siehe: Stolz-Gmeiner, *Theoretische Arithmetik* 85; speziell über periodische systematische Brüche: Weber-Wellstein, *Enzykl. d. Elem.-Math.*, 2. Aufl. (1906), **1**, 254. Über andere Darstellungen reeller Zahlen s. unter *„Kettenbrüche"*.

Die Einführung der irrationalen Zahlen durch Schnitte im Gebiete der rationalen Zahlen stammt von R. Dedekind, *Stetigkeit und irrationale Zahlen*, Braunschweig 1872. Sie wurde seither vielfach wiedergegeben. Wir nennen von neueren Büchern: C. Jordan, *Cours d'analyse*, 2. éd., Paris 1893, **1**, 1 ff.. J. Tannery, *Introduction à la théorie des fonctions d'une variable*,

2. éd., Paris 1904, 1 ff., H. Weber, *Lehrbuch der Algebra,* 2. Aufl., Braunschweig 1898, **1**, 5 ff., Weber-Wellstein, *Enzykl. d. Elem.-Math.*, 2. Aufl., **1**, 76; vgl. auch M. Pasch, *Einleitung in die Differential- und Integralrechnung,* Leipzig 1882, 1 ff.

Mit Hilfe von Folgen rationaler Zahlen wurden die Irrationalzahlen eingeführt von G. Cantor (*Math. Ann.* **5**, 123 (1872), zuerst publiziert von G. Heine, *Journ. f. Math.* **74**, 172 (1872)) und Ch. Méray (*Rev. des soc. sav. (sc. math.)* (2) **4**, 284 (1869), *Nouveau précis d'analyse infinitésimale,* Paris 1872, 1 ff.). Ausführlich findet sich diese Theorie bei Stolz-Gmeiner, *Theoretische Arithmetik* 138 ff. Siehe auch P. Bachmann, *Vorl. über die Natur der Irrationalzahlen,* Leipzig 1892, 6 ff.

Eine andere Theorie wurde von K. Weierstraß in seinen Vorlesungen vorgetragen. Zuerst publiziert von H. Kossak, *Programm des Werderschen Gymnasiums,* Berlin 1872. Später S. Pincherle, *Giorn. di mat.* **18**, 178 und O. Biermann, *Theorie der analytischen Funktionen,* Leipzig 1887, 19.

Eine vergleichende Besprechung dieser drei Theorien findet man bei G. Cantor, *Math. Ann.* **21**, 564, H. Burkhardt, *Vierteljahrschr. Zürich* **46**, 179, O. Perron, *Math.-Ver.* **16**, 142. P. du Bois-Reymond hat diese arithmetischen Theorien in seiner *Allgemeinen Funktionentheorie* (Tübingen 1882) bekämpft. Eine eingehende Darstellung der historischen Entwickelung des Begriffes der irrationalen Zahlen gibt A. Pringsheim, *Enzykl.* I A 2; vgl. auch die Übersetzung dieses Artikels (J. Molk) in der französischen Ausgabe.

Die reellen Zahlen lassen sich geometrisch interpretieren durch die Punkte einer geraden Linie, auf der ein Nullpunkt und eine Einheitsstrecke willkürlich angenommen werden (*Zahlenlinie*).

Diese Zuordnung zwischen den reellen Zahlen und den Punkten (oder Segmenten) einer Geraden ist ein spezieller Fall der Zuordnung der reellen Zahlen zu den Größen eines sogenannten *stetigen Größensystemes.* Siehe: R. Bettazzi, *Teoria delle grandezze*, Pisa 1890, Stolz-Gmeiner, *l. c.* 99, O. Hölder, *Leipz. Ber.* **53**, 1 (1901), E. V. Huntington, *Am. Trans.* **3**, 264 (1902).

Im Vorstehenden wurde gezeigt, wie das System der reellen Zahlen durch sukzessive Erweiterung des Systemes der natürlichen Zahlen gewonnen wird. Man kann auch von vornherein das gesamte Gebiet der reellen Zahlen durch geeignete Axiome einführen: D. Hilbert, *Math.-Ver.* **8**, 180, und *Grundlagen der Geometrie,* 2. Aufl., Leipzig 1903, 24.

§ 3. Die gemeinen komplexen Zahlen.

Wir erweitern das Größensystem der reellen Zahlen durch Hinzufügung aller *Paare reeller Zahlen* (a, a'). Zwei solche Paare (a, a') und (b, b') sollen (anders als es in § 1 geschah) dann und nur dann einander gleich heißen, wenn $a = b$ und $a' = b'$. Das Paar $(a, 0)$ heiße gleich der reellen Zahl a.

Wir definieren eine *Addition* der Zahlenpaare durch:

$$(a, a') + (b, b') = (a + b, a' + b');$$

diese Addition ist *assoziativ, kommutativ,* und läßt eine eindeutige Umkehrung, die *Subtraktion,* zu. Die *Multiplikation* definieren wir durch:

$$(a, a') \cdot (b, b') = (ab - a'b', ab' + a'b);$$

sie ist *assoziativ, kommutativ,* und in Verbindung mit der Addition *distributiv*; ihre Umkehrung, die *Division,* ist immer eindeutig ausführbar, außer durch das Zahlenpaar $(0, 0)$ (das nach der oben gegebenen Definition gleich der Zahl 0 ist).

Ein Produkt kann nur verschwinden, wenn mindestens einer seiner Faktoren verschwindet.

Ein Paar (a, a') wird mit einer reellen Zahl b — oder mit dem der Zahl b gleichen Paare $(b, 0)$ — *multipliziert, indem man jede seiner Zahlen mit b multipliziert.*

Führt man für das Paar $(0, 1)$ die Bezeichnung i ein, so hat man:

$$(a, a') = (a, 0) + (0, a') = a + ia'.$$

Der Ausdruck $a + ia'$ heißt eine *gemeine komplexe Zahl; a* heißt ihr *reeller, ia'* ihr *imaginärer* Teil; i heißt die *imaginäre Einheit.* Eine Zahl, deren reeller Teil Null ist, heißt *rein imaginär.*

Das Quadrat der imaginären Einheit ist -1; denn:

$$i^2 = (0, 1) \cdot (0, 1) = (-1, 0) = -1.$$

Die oben angeführten Sätze über Zahlenpaare nehmen nun, für komplexe Zahlen ausgesprochen, die Form an:

Zwei komplexe Zahlen sind dann und nur dann einander gleich, wenn ihre reellen und imaginären Teile für sich einander gleich sind.

Eine komplexe Zahl ist dann und nur dann gleich Null, wenn sowohl ihr reeller als auch ihr imaginärer Teil Null ist.

$$(a + ia') + (b + ib') = (a + b) + i(a' + b');$$
$$(a + ia') - (b + ib') = (a - b) + i(a' - b');$$
$$(a + ia') \cdot (b + ib') = (ab - a'b') + i(ab' + a'b);$$
$$(a + ia') : (b + ib') = \frac{ab + a'b'}{b^2 + b'^2} + i\frac{a'b - ab'}{b^2 + b'^2}.$$

In der letzten dieser Gleichungen ist vorausgesetzt, daß b und b' nicht beide Null sind.

Sei $a + ia'$ eine von Null verschiedene komplexe Zahl. Setzt man $\varrho = \sqrt[+]{a^2 + a'^2}$ und wird φ bestimmt aus[1]):

$$\cos\varphi = \frac{a}{\sqrt[+]{a^2 + a'^2}}; \quad \sin\varphi = \frac{a'}{\sqrt[+]{a^2 + a'^2}},$$

so erhält man die *trigonometrische Form* der gemeinen komplexen Zahlen:

$$a + ia' = \varrho(\cos\varphi + i\sin\varphi).$$

Die reelle positive Zahl ϱ heißt der *absolute Betrag* (*Modul*) von $a + ia'$ und wird bezeichnet mit $|a + ia'|$; der Winkel φ heißt das *Argument* (Phase, Amplitude) von $a + ia'$.

Sind α und β komplexe Zahlen, so ist:

$$|\alpha| - |\beta| \leqq |\alpha + \beta| \leqq |\alpha| + |\beta|.$$

Der absolute Betrag des Produktes (des Quotienten) zweier komplexer Zahlen ist das Produkt (der Quotient) ihrer absoluten Beträge. Das Argument des Produktes (Quotienten) ist die Summe (Differenz) der Argumente.

Das liefert die *Moivresche Formel* für die n^{te} Potenz einer komplexen Zahl

$$[\varrho(\cos\varphi + i\sin\varphi)]^n = \varrho^n(\cos n\varphi + i\sin n\varphi),$$

aus der man durch Entwickelung der linken Seite und Trennung des Reellen vom Imaginären Formeln für $\cos n\varphi$ und $\sin n\varphi$ erhält.

Die Zahl $a - ib$ heißt die zu $a + ib$ *konjugierte* Zahl. Sie hat denselben absoluten Betrag. Die Summe sowie das Produkt zweier konjugiert-komplexer Zahlen ist reell, und zwar ist ihr Produkt gleich dem Quadrat ihres absoluten Betrages.

---

1) Diese Gleichungen bestimmen φ nur bis auf Vielfache von $2\cdot$ Der den Ungleichungen $-\pi < \varphi \leqq +\pi$ genügende Wert von φ wi als *Hauptwert* bezeichnet.

Sei $\alpha_n = a_n + ia'_n$ $(n = 1, 2, \ldots)$ eine *Folge komplexer Zahlen.* Sie heißt k .vergent, wenn zu jeder positiven Zahl ε ein Wert N des Index n gehört, derart, daß für alle Indices n' und n'', die größer als N sind, die Ungleichung besteht:

$$|\alpha_{n'} - \alpha_{n''}| < \varepsilon.$$

Eine Folge komplexer Zahlen, die nicht konvergent ist, heißt *divergent.*

Die komplexe Zahl $\alpha = a + ia'$ heißt *Grenze* der Folge (α_n), in Zeichen:

$$\lim_{n=\infty} \alpha_n = \alpha,$$

wenn zu jedem positiven ε ein N gehört, derart, daß für jeden Index n, der größer als N ist, die Ungleichung besteht:

$$|\alpha - \alpha_n| < \varepsilon.$$

Damit die Folge komplexer Zahlen $(a_n + ia'_n)$ *konvergent sei, ist notwendig und hinreichend, daß die beiden Folgen reeller Zahlen* (a_n) *und* (a'_n) *konvergent sind.*

Damit die Folge komplexer Zahlen $(a_n + ia'_n)$ *die Zahl* $a + ia'$ *zur Grenze habe, ist notwendig und hinreichend, daß die Folge* (a_n) *die Zahl* a, *die Folge* (a'_n) *die Zahl* a' *zur Grenze hat.*

Die komplexen Zahlen lassen sich geometrisch durch die Punkte einer Ebene darstellen, indem man, unter Zugrundelegung eines rechtwinkligen Koordinatensystems, der Zahl $a + ia'$ den Punkt A mit der Abszisse a und der Ordinate a' zuordnet. Dann ist der absolute Betrag von $a + ia'$ gleich dem Abstande des Punktes A vom Koordinatenursprung O, und das Argument φ von $a + ia'$ ist gleich dem Winkel, den die Richtung OA mit der positiven x-Achse bildet.

Den arithmetischen Grundoperationen an komplexen Zahlen entsprechen folgende geometrische Konstruktionen:

Seien A und A' die zwei komplexen Zahlen α und α' entsprechenden Punkte. *Trägt man von A aus eine der Strecke OA' gleiche und gleich gerichtete Strecke auf, so erhält man den der Summe* $\alpha + \alpha'$ *entsprechenden Punkt. Trägt man von A aus eine der Strecke OA' gleiche, aber entgegengesetzt gerichtete Strecke auf, so erhält man den der Differenz* $\alpha - \alpha'$ *entsprechenden Punkt.*

Um den dem *Produkte* $\alpha \cdot \alpha'$ entsprechenden Punkt zu konstruieren, nehme man auf der Achse der reellen Zahlen den Punkt der Abszisse 1 und errichte über der Strecke OA' das

Dreieck $OB1'$ ähnlich zu $OA1$, und zwar so, daß die Winkel BOA' und $AO1$ gleichen Sinn haben. Der Punkt B ist der gesuchte.

Um den dem *Quotienten* $\frac{\alpha}{\alpha'}$ entsprechenden Punkt zu konstruieren, errichte man über OA das Dreieck OBA ähnlich zum Dreieck $O1A'$, und zwar so, daß die Winkel BOA und $A'O1$ entgegengesetzten Sinn haben. Der Punkt B ist der gesuchte.

Die komplexen Zahlen sind in der Mathematik allgemein anerkannt seit K. F. Gauss (vgl. *Demonstratio nova theorematis etc.* (1799), *Werke* **3**, 3 u. *Theoria residuorum biquadraticorum*(1831), *Werke* **2**,171). Die hier skizzierte Einführung der komplexen Zahlen durch Paare reeller Zahlen rührt von W. R. Hamilton her (*Dublin Transact.* **17**, 393 (1837), *Lectures on Quaternions* 1853, Vorrede). Die geometrische Deutung der komplexen Zahlen durch Punkte einer Ebene geht zurück auf C. Wessel (1799), (die betreffende Arbeit wurde neu herausgegeben unter dem Titel: *Essai sur la représentation analytique de la direction,* Kopenhagen 1897) und R. Argand (*Essai sur une manière de représenter les quantités imaginaires,* Paris 1806). Analog ist die Deutung durch die Vektoren einer Ebene in der Äquipollenzenrechnung, mit der sich G. Bellavitis (seit 1833) viel beschäftigte. Es sei auf C. A. Laisant, *Théorie et applications des équipollences*, Paris 1887, verwiesen.

§ 4. Die höheren komplexen Zahlen.

Analog wie im vorigen Paragraphen die Zahlenpaare betrachtet wurden, betrachten wir hier die „n-tupel" reeller Zahlen $(a_1, a_2, \ldots, a_n)$. Wir definieren:

Zwei n-tupel reeller Zahlen $(a_1, a_2, \ldots, a_n)$ und $(b_1, b_2, \ldots, b_n)$ heißen *einander gleich,* wenn $a_1 = b_1$, $a_2 = b_2$, $\ldots$, $a_n = b_n$.

Das n-tupel $(a_1, a_2, \ldots, a_n)$ heißt *gleich Null,* wenn $a_1 = a_2 = \cdots = a_n = 0$ ist.

Die *Addition* der n-tupel wird definiert durch:

$$(a_1, a_2, \ldots, a_n) + (b_1, b_2, \ldots, b_n) = (a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n).$$

Multiplikation eines n-tupels mit einer reellen Zahl b wird definiert durch:

$$b\,(a_1, a_2, \ldots, a_n) = (a_1, a_2, \ldots, a_n)\,b = (a_1 b, a_2 b, \ldots, a_n b).$$

Führt man nun für die speziellen n-tupel:

$$(1, 0, \ldots, 0), (0, 1, 0, \ldots, 0) \ldots (0, 0, \ldots, 0, 1)$$

die Bezeichnungen $e_1, e_2, \ldots, e_n$ ein, so hat man:

$$(a_1, a_2, \ldots, a_n) = a_1 e_1 + a_2 e_2 + \cdots + a_n e_n.$$

Ein solcher Ausdruck heißt *eine komplexe Zahl mit den n Einheiten* $e_1, e_2, \ldots, e_n$.

Wir haben also die Sätze:

Zwei komplexe Zahlen mit den n Einheiten $e_1, e_2, \ldots, e_n$ *sind einander gleich, wenn die Koeffizienten entsprechender Einheiten einander gleich sind.*

Eine solche Zahl ist gleich Null, wenn die Koeffizienten sämtlicher Einheiten Null sind.

Zwei komplexe Zahlen mit den n Einheiten $e_1, e_2, \ldots, e_n$ *werden addiert, indem man die Koeffizienten entsprechender Einheiten addiert.*

Diese Addition ist *assoziativ* und *kommutativ*; sie läßt eine eindeutige Umkehrung, die *Subtraktion,* zu.

Um die *Multiplikation* unserer komplexen Zahlen zu definieren, definiert man zuerst die „Einheitsprodukte" $e_i e_k$ als Zahlen des Systemes durch die Formeln:

$$e_i e_k = \lambda_{ik}^{(1)} e_1 + \lambda_{ik}^{(2)} e_2 + \cdots + \lambda_{ik}^{(n)} e_n,$$

wo die $\lambda_{ik}^{(r)}$ reelle Zahlen bedeuten; sodann definiert man das Produkt der beiden Zahlen

$$\sum_{i=1}^{n} a_i e_i \quad \text{und} \quad \sum_{k=1}^{n} b_k e_k$$

durch:

$$\sum_{i=1}^{n} a_i e_i \cdot \sum_{k=1}^{n} b_k e_k = \sum_{i,k=1}^{n} a_i b_k e_i e_k.$$

Diese Multiplikation ist in Verbindung mit der Addition *distributiv.* Damit sie *assoziativ* sei, ist notwendig und hinreichend, daß die $\lambda_{ik}^{(r)}$ den Gleichungen genügen:

$$\sum_{v=1}^{n} \lambda_{ik}^{(v)} \lambda_{vm}^{(r)} = \sum_{v=1}^{n} \lambda_{km}^{(v)} \lambda_{iv}^{(r)} \qquad (i, k, m, r = 1, 2, \ldots, n).$$

Damit die Multiplikation kommutativ sei, ist notwendig und hinreichend, daß $e_i e_k = e_k e_i$ ist, d. h. daß $\lambda_{ik}^{(r)} = \lambda_{ki}^{(r)}$ ist.

Wählt man aus unserem System komplexer Zahlen n Zahlen

$$(1) \qquad i_k = \alpha_{1k} e_1 + \alpha_{2k} e_2 + \cdots + \alpha_{nk} e_n \qquad (k = 1, 2, \ldots, n)$$

so aus, daß die aus den reellen Zahlen α_{ik} gebildete Deter-

minante nicht verschwindet, so läßt sich jede Zahl

$$a_1 e_1 + a_2 e_2 + \cdots + a_n e_n$$

unseres Systemes auch in der Form schreiben:

$$b_1 i_1 + b_2 i_2 + \cdots + b_n i_n. \tag{2}$$

Man sagt: *das System komplexer Zahlen (2) mit den n Einheiten $i_1, i_2, \ldots, i_n$ geht aus unserem ursprünglichen Systeme durch die lineare Transformation (1) hervor.* Umgekehrt geht unser ursprüngliches System aus dem Systeme der Zahlen (2) durch die zur Transformation (1) reziproke Transformation hervor.

Unter allen komplexen Zahlen mit zwei oder mehr Einheiten sind die gemeinen komplexen Zahlen und die daraus durch lineare Transformation entstehenden die einzigen, die eine assoziative, kommutative und in Verbindung mit der Addition distributive Multiplikation zulassen, die so beschaffen ist, daß ein Produkt nur dann verschwindet, wenn mindestens ein Faktor verschwindet (zuerst bewiesen von K. Weierstraß in seinen Vorlesungen (1863); vgl. E. Kossak, *Progr. d. Werderschen Gymn. Berlin* 1872, 23 ff., G. Pincherle, *Giorn. di mat.* **18**, 203 ff.).

Unter *Quaternionen* versteht man nach W. R. Hamilton (*Lectures on Quaternions*, Dublin 1853) komplexe Zahlen mit vier Einheiten:

$$a_0 + a_1 i_1 + a_2 i_2 + a_3 i_3.$$

Die erste Einheit ist hier, wie bei den gemeinen komplexen Zahlen, der reellen Einheit gleichgesetzt. Für die Produkte der übrigen Einheiten gelten die Formeln:

$$\begin{aligned} i_1^2 &= -1 & i_2 i_3 &= -i_3 i_2 = i_1, \\ i_2^2 &= -1 & i_3 i_1 &= -i_1 i_3 = i_2, \\ i_3^2 &= -1 & i_1 i_2 &= -i_2 i_1 = i_3. \end{aligned}$$

Die hierdurch definierte Multiplikation der Quaternionen ist assoziativ und in Verbindung mit der Addition distributiv, aber nicht kommutativ. Ein Produkt kann nur verschwinden, wenn mindestens ein Faktor verschwindet. Es gibt zu dieser Multiplikation zwei Umkehrungen (*Divisionen*), eine vordere und eine hintere; d. h. bedeuten α und β zwei Quaternionen, und ist α nicht Null, so hat jede der beiden Gleichungen: $\xi \cdot \alpha = \beta$ und $\alpha \cdot \eta = \beta$ eine und nur eine Quaternion zur Lösung.

Unter allen komplexen Zahlen mit zwei oder mehr Einheiten gibt es außer den gemeinen komplexen Zahlen und den Quater-

nionen (sowie den daraus durch lineare Transformation entstehenden Systemen) keine, die eine assoziative und in Verbindung mit der Addition distributive Multiplikation zulassen, die so beschaffen ist, daß ein Produkt nur dann verschwindet, wenn wenigstens ein Faktor verschwindet (*Satz von G. Frobenius, Journ. f. Math.* **84**, 59 (1878)).

Wegen der Anwendungen der Quaternionen in der Vektorenrechnung verweisen wir auf W. R. Hamilton, *Elements of Quaternions.* 2. ed., 2 Bde., London 1899, 1901 (deutsch von P. Glan, Leipzig 1882/84) und P. G. Tait, *An elementary treatise on Quaternions,* 3. ed., Cambridge 1890 (deutsch von G. v. Scherff, Leipzig 1880).

Die in diesem und dem vorhergehenden Paragraphen skizzierten Theorien findet man ausführlich bei Stolz-Gmeiner, *Theoretische Arithmetik* 277 ff. Näheres über komplexe Zahlen sowie die darauf bezügliche Literatur bei E. Study, *Enzykl.* **I A 4**; vgl. auch Kap. II, § 7.

Man kann auch komplexe Zahlen mit unendlich vielen Einheiten betrachten. Sie kommen zur Verwendung in der Theorie der sogenannten *nichtarchimedischen Größensysteme.* Siehe G. Veronese, *Fondamenti di geometria*, Padova 1891, deutsch von A. Schepp, Leipzig 1894 (vgl. besonders die Anm. auf S. 139 der deutschen Ausgabe) und H. Hahn, *Über die nichtarchimedischen Größensysteme, Wien. Ber.* **116**, 601 (1907), K. Th. Vahlen, *Math.-Ver.* **16**, 409 (1907).

§ 5. Grundbegriffe der Mengenlehre. Die Mächtigkeiten der Mengen.

Unter einer *Menge* wird die Zusammenfassung wohldefinierter und unterscheidbarer Dinge (Elemente) zu einem Ganzen verstanden.[1])

Eine Menge N, deren sämtliche Elemente auch der Menge M angehören, heißt *Teilmenge* von M. Gibt es in M wenigstens ein Element, das nicht zu N gehört, so heißt N eine *echte* Teilmenge von M.

Zwei Mengen M und N heißen *von gleicher Mächtigkeit* oder *äquivalent* (in Zeichen $M \sim N$), wenn sich ihre Elemente

1) Gegen diese von G. Cantor herrührende Definition (*Math. Ann.* **46**, 481, vgl. auch *Math. Ann.* **20**, 114) bestehen gewisse Bedenken, vgl. J. Mollerup, *Math. Ann.* **64**, 231 und die am Schlusse von § 7 angegebene Literatur.

eineindeutig aufeinander beziehen lassen. Sind zwei Mengen zu ein und derselben dritten Menge äquivalent, so sind sie auch untereinander äquivalent.

Die Gesamtheit aller untereinander äquivalenter Mengen definiert eine bestimmte *Mächtigkeit.* Man bezeichnet häufig die Mächtigkeit einer Menge mit dem entsprechenden deutschen Buchstaben (z. B. die Mächtigkeit von M mit $\mathfrak{m}$).

Eine Menge heißt *endlich,* wenn sie keiner ihrer echten Teilmengen äquivalent ist; eine Menge, die nicht endlich ist, heißt *unendlich* oder *transfinit.*

Jede endliche Menge ist äquivalent einem und nur einem Segmente der Reihe der natürlichen Zahlen. Dabei ist unter einem Segmente der Zahlenreihe die Menge aller jener natürlichen Zahlen verstanden, die eine gewisse natürliche Zahl n nicht übersteigen. Durch diese Zahl n ist die Mächtigkeit unserer Menge vollständig charakterisiert. *Zur Bezeichnung der Mächtigkeiten endlicher Mengen dienen daher die natürlichen Zahlen;* in diesem Sinne verwendet heißen die natürlichen Zahlen *Kardinalzahlen.*

Ist die Menge M äquivalent mit einer Teilmenge von N und die Menge N äquivalent mit einer Teilmenge von M, so sind M und N äquivalent (*Cantors Äquivalenzsatz;* Beweise u. a. von F. Bernstein in É. Borel, *Leçons sur la théorie des fonctions,* Paris (1898), 102, J. König, *C. R.* **143**, 110 (1906), G. Peano, *Rend. Pal.* **31**, 360 (1906)).

Ist die Menge M äquivalent mit einer Teilmenge von N, gibt es aber in M keine zu N äquivalente Teilmenge, so heißt die Mächtigkeit von N *größer* als die von M; in Zeichen: $\mathfrak{n} > \mathfrak{m}$; im entgegengesetzten Falle hat man $\mathfrak{n} < \mathfrak{m}$.

Es ist bisher nicht einwandfrei bewiesen, daß, wenn die beiden Mengen M und N nicht äquivalent sind (also ihre Mächtigkeiten $\mathfrak{m}$ und $\mathfrak{n}$ nicht gleich sind), notwendig einer der beiden Fälle $\mathfrak{n} > \mathfrak{m}$ oder $\mathfrak{n} < \mathfrak{m}$ eintreten muß.

Sei irgendeine (endliche oder unendliche) Menge von Mengen M_a, M_b, ... gegeben, von denen keine zwei ein gemeinsames Element besitzen mögen. Die Menge, welche aus allen in diesen Mengen enthaltenen Elementen besteht, heißt die *Vereinigungsmenge* dieser Mengen. Ihre Mächtigkeit heißt die *Summe* der Mächtigkeiten der einzelnen Mengen.[1]) Diese Addition

1) Speziell erhält man also die Summe $\mathfrak{m} + \mathfrak{n}$ zweier Mächtigkeiten, indem man die Vereinigungsmenge einer Menge M der Mächtigkeit $\mathfrak{m}$ und einer Menge N der Mächtigkeit $\mathfrak{n}$ bildet.

ist *assoziativ* und *kommutativ*. Es gilt die Ungleichung

$$\mathfrak{m} + \mathfrak{n} \geqq \mathfrak{m};$$

und aus $\mathfrak{n} > \mathfrak{n}'$ folgt

$$\mathfrak{m} + \mathfrak{n} \geqq \mathfrak{m} + \mathfrak{n}'.$$

Man greife aus jeder der Mengen M_a, M_b, ... je ein Element heraus: m_a, m_b, ... Die Menge aller so erhältlichen Elementkombinationen $(m_a, m_b, \ldots)$ heißt die *Verbindungsmenge* dieser Mengen. Ihre Mächtigkeit heißt das *Produkt* der Mächtigkeiten der einzelnen Mengen. Diese Multiplikation ist *assoziativ, kommutativ* und in Verbindung mit der Addition *distributiv*. Es ist

$$\mathfrak{m} \cdot \mathfrak{n} \geqq \mathfrak{m}$$

und aus $\mathfrak{n} > \mathfrak{n}'$ folgt

$$\mathfrak{m} \cdot \mathfrak{n} \geqq \mathfrak{m} \cdot \mathfrak{n}'.$$

Das Produkt $\mathfrak{m} \cdot \mathfrak{n}$ *ist gleich einer Summe von lauter einander gleichen Summanden* $\mathfrak{m}$, *wenn diese Summanden eine Menge der Mächtigkeit* $\mathfrak{n}$ *bilden.*

Man ordne jedem Element der Menge N ein beliebiges Element der Menge M zu; die Menge aller so erhältlichen Zuordnungen heißt die *Belegungsmenge von N mit M.* Ihre Mächtigkeit wird bezeichnet mit $\mathfrak{m}^{\mathfrak{n}}$. Für diese Potenzierung gelten die Gesetze:

$$\mathfrak{m}^{\mathfrak{p}} \cdot \mathfrak{n}^{\mathfrak{p}} = (\mathfrak{m} \cdot \mathfrak{n})^{\mathfrak{p}}; \quad \mathfrak{m}^{\mathfrak{n}} \cdot \mathfrak{m}^{\mathfrak{p}} = \mathfrak{m}^{\mathfrak{n}+\mathfrak{p}}; \quad (\mathfrak{m}^{\mathfrak{n}})^{\mathfrak{p}} = \mathfrak{m}^{\mathfrak{n} \cdot \mathfrak{p}}$$

Ferner ist $\mathfrak{m}^{\mathfrak{n}} \geqq \mathfrak{m}$ und aus $\mathfrak{n} > \mathfrak{n}'$ folgt $\mathfrak{m}^{\mathfrak{n}} \geqq \mathfrak{m}^{\mathfrak{n}'}$.

Die Potenz $\mathfrak{m}^{\mathfrak{n}}$ *ist gleich einem Produkte lauter einander gleicher Faktoren* $\mathfrak{m}$, *wenn diese Faktoren eine Menge der Mächtigkeit* $\mathfrak{n}$ *bilden.*

Es bestehen für $\mathfrak{m} > 1$, $\mathfrak{n} > 1$ die Ungleichungen:

$$\mathfrak{m} + \mathfrak{n} \leqq \mathfrak{m} \cdot \mathfrak{n} \leqq \mathfrak{m}^{\mathfrak{n}}.$$

Ferner ist für $\mathfrak{m} > 1$ stets $\mathfrak{m}^{\mathfrak{m}} > \mathfrak{m}$; *daher kann es keine größte Mächtigkeit geben.*

Die Menge aller Teilmengen einer transfiniten Menge der Mächtigkeit $\mathfrak{m}$ *hat die Mächtigkeit* $2^{\mathfrak{m}}$ *und diese Mächtigkeit ist größer als* $\mathfrak{m}$.

Eine Menge, deren Elemente sich eineindeutig den natürlichen Zahlen zuordnen lassen, heißt *abzählbar*. Ihre Mächtigkeit wird bezeichnet mit $\aleph_0$ (oft auch mit $\mathfrak{a}$).

Es ist $\aleph_0 + \aleph_0 = \aleph_0$ *und* $\aleph_0 \cdot \aleph_0 = \aleph_0$.

Eine unendliche Teilmenge einer abzählbaren Menge ist abzählbar.

Jede unendliche Menge enthält eine abzählbare Teilmenge (d. h.: $\aleph_0$ ist die kleinste transfinite Mächtigkeit).

Eine unendliche Menge ändert durch Hinzufügung einer endlichen oder abzählbaren Menge ihre Mächtigkeit nicht.

Abzählbare Mengen sind: die Menge aller (positiven und negativen) ganzen Zahlen, die Menge aller rationalen Zahlen, die Menge aller algebraischen Zahlen (dabei ist als *algebraisch* jede Zahl bezeichnet, die einer algebraischen Gleichung mit ganzzahligen Koeffizienten genügt). Jede unendliche Menge sich nicht überdeckender Intervalle einer Geraden (oder sich nicht überdeckender Gebiete einer Ebene, eines n dimensionalen Raumes) ist abzählbar.

Man sagt von einer Menge, *sie hat die Mächtigkeit des Kontinuums*, wenn sich ihre Elemente eineindeutig den reellen Zahlen zuordnen lassen. Auch die Menge aller reellen Zahlen, die zwischen zwei gegebenen reellen Zahlen liegen (oder, was dasselbe ist, die Menge aller Punkte eines beliebigen Intervalles), hat die Mächtigkeit des Kontinuums.

Die Mächtigkeit des Kontinuums wird mit $\mathfrak{c}$ bezeichnet; *es ist* $\mathfrak{c} > \aleph_0$ *und es besteht die Gleichung:*

$$\mathfrak{c} = \mathfrak{e}^{\aleph_0} = \aleph_0^{\aleph_0},$$

in der $\mathfrak{e}$ eine beliebige endliche Kardinalzahl ($\geqq 2$) bedeutet.

Da hiernach die Menge aller reellen Zahlen größere Mächtigkeit hat als die aller algebraischen Zahlen, folgt die Existenz nicht algebraischer reeller Zahlen; es sind die sogenannten *transzendenten* Zahlen.

Es ist $\mathfrak{c}^2 = \mathfrak{c}$; allgemein $\mathfrak{c}^{\mathfrak{e}} = \mathfrak{c}$, wenn $\mathfrak{e}$ eine endliche Kardinalzahl; *somit lassen sich die Punkte einer Ebene (eines Raumes von beliebig viel Dimensionen) eineindeutig den Punkten einer Geraden zuordnen.* Es besteht auch die Gleichung: $\mathfrak{c}^{\aleph_0} = \mathfrak{c}$, die man auch so aussprechen kann: die Punkte eines Raumes von abzählbar unendlich vielen Dimensionen lassen sich eineindeutig den Punkten einer Geraden zuordnen.

§ 6. Die geordneten Mengen. Die Ordnungstypen.

Eine Menge heißt *einfach geordnet*, wenn — zufolge irgendeiner Festsetzung — von je zweien ihrer Elemente (m und m') das eine (m) als das vorhergehende, das andere (m') als das nachfolgende definiert ist (in Zeichen: $m \prec m'$ oder $m' \succ m$). Dabei hat diese Festsetzung der Rangordnung der Forderung zu genügen: wenn $m \prec m'$ und $m' \prec m''$, so ist stets auch $m \prec m''$.

Steht ein Element m_0 der einfach geordneten Menge M zu jedem anderen Elemente m dieser Menge in der Beziehung $m_0 \prec m$ ($m_0 \succ m$), so heißt es das *erste* (*letzte*) Element von M. Ist $m \succ m'$ und $m \prec m''$, so sagt man: m liegt zwischen m' und m''.

Zwei einfach geordnete Mengen M und N heißen *ähnlich* oder vom gleichen *Ordnungstypus* (in Zeichen $M \simeq N$), wenn sich ihre Elemente so eineindeutig zuordnen lassen, daß die Rangordnung erhalten bleibt, d. h. so, daß, wenn m_1 und m_2 irgendzwei Elemente von M, n_1 und n_2 die entsprechenden Elemente von N sind, aus $m_1 \prec m_2$ immer folgt $n_1 \prec n_2$. Sind zwei Mengen derselben dritten Menge ähnlich, so sind sie auch untereinander ähnlich.

Die Gesamtheit aller untereinander ähnlichen Mengen definiert einen bestimmten *Ordnungstypus*. Man bezeichnet häufig den Ordnungstypus einer Menge mit dem entsprechenden griechischen Buchstaben (z. B. den Ordnungstypus von M mit μ).

Alle Mengen vom gleichen Ordnungstypus μ haben die gleiche Mächtigkeit, die kurz die Mächtigkeit von μ genannt wird; die Umkehrung gilt nur für endliche Mengen:

Zwei einfach geordnete endliche Mengen von gleicher Mächtigkeit haben auch gleichen Ordnungstypus.

Zur Bezeichnung der Ordnungstypen endlicher Mengen können daher, wie zur Bezeichnung ihrer Mächtigkeiten, die natürlichen Zahlen verwendet werden; in diesem Sinne verwendet heißen sie *Ordinalzahlen*.

Man definiert die *Addition* der Ordnungstypen in folgender Weise: Seien zwei einfach geordnete Mengen M und N ohne gemeinsames Element gegeben; ihre Ordnungstypen seien μ und ν. Wir bilden ihre Vereinigungsmenge und machen sie zu einer einfachen geordneten Menge durch die Festsetzung: die Elemente von M untereinander, ebenso die von N, mögen ihre frühere Rangordnung beibehalten, jedes Element von M aber gehe jedem Element von N voran. Der Ordnungstypus der so geordneten Vereinigungsmenge wird mit $\mu + \nu$ bezeichnet. (Ähnlich wie für die Mächtigkeiten in § 5, kann man auch hier, statt sich auf die Definition der Summe von *zwei* Ordnungstypen zu beschränken, sofort die Summe der Ordnungstypen von Mengen definieren, die selbst eine einfach geordnete Menge bilden.)

Es ist $(\mu + \nu) + \pi = \mu + (\nu + \pi)$, aber im allgemeinen $\mu + \nu \neq \nu + \mu$.

Um die *Multiplikation* der Ordnungstypen zu definieren,

bildet man die Verbindungsmenge von M und N, d. h. die Menge der Paare (m, n) und ordnet sie durch die Festsetzung: $(m, n) \prec (m', n')$, wenn $n \prec n'$, und im Falle $n = n'$, wenn $m \prec m'$. Der Ordnungstypus der so geordneten Verbindungsmenge wird mit $\mu \cdot \nu$ bezeichnet.

Es ist $(\mu \cdot \nu) \cdot \pi = \mu \cdot (\nu \cdot \pi)$ und $\mu \cdot (\nu + \pi) = \mu \cdot \nu + \mu \cdot \pi$, aber im allgemeinen $\mu \cdot \nu \neq \nu \cdot \mu$ und $(\mu + \nu) \cdot \pi \neq \mu \cdot \pi + \nu \cdot \pi$.

Eine Definition der Potenz μ^ν findet man bei F. Hausdorff, *Leipz. Ber.* **58**, 106 (1906), G. Hessenberg, *Math.-Ver.*, **16**, 130 (1907).

Eine Folge $m_1, m_2, \ldots, m_n, \ldots$ von Elementen von M heißt eine *aufsteigende Fundamentalreihe,* wenn

$$m_1 \prec m_2 \prec \cdots \prec m_n \prec \cdots;$$

sie heißt eine *absteigende Fundamentalreihe,* wenn

$$m_1 \succ m_2 \succ \cdots \succ m_n \succ \cdots.$$

Ein Element m von M heißt *Grenzelement* der aufsteigenden Fundamentalreihe $m_1, m_2, \ldots, m_n, \ldots$, wenn 1. $m \succ m_n$ für jedes n und es 2. in M kein Element m' gibt, so daß ebenfalls $m' \succ m_n$ für jedes n, und außerdem $m' \prec m$ ist. Analog wird ein Grenzelement einer absteigenden Fundamentalreihe definiert.

Der Ordnungstypus μ einer Menge M heißt *abgeschlossen,* wenn jede in M enthaltene Fundamentalreihe ein Grenzelement besitzt. Er heißt *in sich dicht,* wenn jedes Element von M Grenzelement einer in M enthaltenen Fundamentalreihe ist. Er heißt *perfekt,* wenn er sowohl abgeschlossen als auch in sich dicht ist. Er heißt *überall dicht,* wenn zwischen irgend zwei Elementen von M stets noch ein Element von M liegt.

Der Ordnungstypus der Menge der natürlichen Zahlen in ihrer natürlichen Reihenfolge ($m \prec n$ wenn $m < n$) wird mit ω bezeichnet; der Ordnungstypus dieser Menge in der umgekehrten Reihenfolge ($m \prec n$ wenn $m > n$) wird mit ω^* bezeichnet. Demnach ist der Ordnungstypus der Menge aller (positiven und negativen) ganzen Zahlen in ihrer natürlichen Reihenfolge: $\omega^* + \omega$.

Der Ordnungstypus der Menge aller echten rationalen Brüche ($0 < r < 1$) in ihrer natürlichen Reihenfolge ($r' \prec r''$ wenn $r' < r''$) wird mit η bezeichnet.

Jede abzählbare Menge ohne erstes und letztes Element von überall dichtem Ordnungstypus hat den Ordnungstypus η.

Der Ordnungstypus der Menge aller reellen Zahlen von 0 bis 1 $(0 \leqq x \leqq 1)$ in ihrer natürlichen Reihenfolge wird mit ϑ bezeichnet.

Jede Menge M von perfektem Ordnungstypus, die eine abzählbare Teilmenge A enthält, derart, daß zwischen irgendzwei Elementen von M stets ein Element von A liegt, hat den Ordnungstypus ϑ.

§ 7. Die wohlgeordneten Mengen.

Eine einfach geordnete Menge heißt *wohlgeordnet,* wenn jede ihrer Teilmengen ein erstes Element enthält.

Jede einfach geordnete endliche Menge ist wohlgeordnet. Auch die Menge der natürlichen Zahlen in ihrer natürlichen Reihenfolge ist wohlgeordnet. Jede Teilmenge einer wohlgeordneten Menge ist wohlgeordnet.

In einer wohlgeordneten Menge gibt es zu jedem Element (wenn es nicht das letzte Element der Menge ist) ein unmittelbar folgendes.

Diejenige Teilmenge einer wohlgeordneten Menge M, die aus sämtlichen dem Element m vorangehenden Elementen von M besteht, heißt *der Abschnitt von m* und wird mit $A(m)$ bezeichnet. Um gewisse Sätze leichter aussprechen zu können, ordnet man auch dem ersten Element m_0 von M einen (fingierten) Abschnitt $A(m_0)$ zu, der kein einziges Element enthält.

Eine wohlgeordnete Menge ist keinem ihrer Abschnitte ähnlich. Ebenso können zwei verschiedene Abschnitte derselben wohlgeordneten Menge einander nicht ähnlich sein.

Bildet man die Menge aller Abschnitte $A(m)$ der wohlgeordneten Menge M (inbegriffen den Abschnitt $A(m_0)$) und ordnet sie durch die Vorschrift: $A(m) \prec A(m')$, wenn $m \prec m'$, so erhält man eine wohlgeordnete Menge, die der Menge M ähnlich ist.

Zwei beliebige wohlgeordnete Mengen sind entweder einander ähnlich, oder es ist eine von beiden einem Abschnitte der anderen ähnlich.

Die Ordnungstypen wohlgeordneter Mengen heißen *Ordinalzahlen,* speziell die der transfiniten wohlgeordneten Mengen heißen *transfinite Ordinalzahlen.* Um gewisse Sätze leichter aussprechen zu können, betrachtet man auch die Null als Ordinalzahl, und zwar als die Ordinalzahl einer Menge, die kein einziges Element enthält, wie etwa der oben eingeführte Abschnitt $A(m_0)$.

Die Ordinalzahlen μ und ν zweier wohlgeordneter Mengen M und N heißen einander gleich, wenn M und N ähnlich sind. Die Ordinalzahl μ heißt kleiner als ν (in Zeichen: $\mu < \nu$), wenn M einem Abschnitt von N ähnlich ist; dementsprechend ist μ größer als ν (in Zeichen: $\mu > \nu$), wenn N einem Abschnitt von M ähnlich ist. Die Null ist kleiner als jede andere Ordinalzahl.

Die Gesamtheit aller Ordinalzahlen, die kleiner sind als eine gegebene Ordinalzahl μ, bildet eine wohlgeordnete Menge, deren Ordinalzahl μ ist.

Die *Addition* und *Multiplikation* der Ordinalzahlen ergibt sich als Spezialfall der im vorigen Paragraphen definierten Addition und Multiplikation der Ordnungstypen.

Zu jeder Ordinalzahl μ gibt es eine nächstgrößere, es ist die Ordinalzahl $\mu + 1$.

Zu jeder unendlichen Menge M von Ordinalzahlen μ, die keine größte Zahl enthält, gibt es eine nächstgrößere Ordinalzahl (d. h. eine Ordinalzahl, die größer ist als alle in M enthaltenen Zahlen, und kleiner als jede andere Ordinalzahl, die ebenfalls größer als alle Zahlen von M ist). Diese Zahl wird der Limes der in M enthaltenen Zahlen genannt und mit $\lim \mu$ bezeichnet; so ist, wenn mit n eine endliche Ordinalzahl bezeichnet wird: $\lim n = \omega$.

Ist die Zahl ν so beschaffen, daß es unter den Zahlen μ, die kleiner als ν sind, keine größte gibt, so heißt ν eine *Grenzzahl* oder *Limeszahl*, und zwar ist

$$\nu = \lim \mu \; (\mu < \nu).$$

Die *Potenzierung* der Ordinalzahlen wird definiert durch die drei Festsetzungen:

$$\mu^0 = 1; \quad \mu^{\nu+1} = \mu^\nu \cdot \mu; \quad \mu^{\lim \nu} = \lim \mu^\nu.$$

Jede Ordinalzahl kann (und zwar nur auf eine einzige Weise) in der Form geschrieben werden:

$$\mu = \omega^{\nu_0} n_0 + \omega^{\nu_1} n_1 + \cdots + \omega^{\nu_k} n_k,$$

wo die rechtsstehende Summe nur eine endliche Anzahl von Gliedern enthält, $\nu_0, \nu_1, \ldots, \nu_k$ irgendwelche Ordinalzahlen, $n_0, n_1, \ldots, n_k$ aber endliche Ordinalzahlen bedeuten (Cantors Normalform).

Es gibt Ordinalzahlen ε, die der Gleichung $\omega^\varepsilon = \varepsilon$ genügen, sie heißen ε-Zahlen.

Die kleinste transfinite Ordinalzahl ist ω; es folgen

$$\omega + 1, \omega + 2, \ldots, \omega + n, \ldots$$

(wo n eine endliche Ordinalzahl). Der Limes der Zahlen $\omega + n$ ist $\omega \cdot 2$; es folgen $\omega \cdot 2 + 1$, $\omega \cdot 2 + 2$, allgemein $\omega \cdot 2 + n$; auf alle diese folgt $\omega \cdot 3$ etc. Allgemein erhält man so Zahlen der Form $\omega \cdot m + n$, wo m und n endliche Zahlen. Der Limes aller dieser Zahlen ist ω^2; indem man so fortfährt, erhält man Zahlen der Form

$$\omega^{m_0} n_0 + \omega^{m_1} n_1 + \cdots + \omega^{m_k} n_k,$$

wo

$$m_0, m_1, \ldots, m_k, n_0, n_1, \ldots, n_k$$

endliche Zahlen sind; der Limes aller dieser Zahlen ist ω^ω. Analog erhält man

$$\omega^{\omega^\omega}, \ldots, \omega^{\omega^{\cdot^{\cdot^{\cdot^{\omega}}}}}, \ldots.$$

Der Limes aller dieser Zahlen ist die erste ε Zahl usw.

Jeder transfiniten Ordinalzahl μ kommt (wie jedem Ordnungstypus, siehe § 6) eine bestimmte Mächtigkeit zu. Die Mächtigkeiten transfiniter wohlgeordneter Mengen werden mit dem Buchstaben $\aleph$ (Alef) bezeichnet, und die verschiedenen Alefs durch Indices unterschieden. Von zwei gegebenen Alefs, die nicht einander gleich sind, ist stets das eine das größere, das andere das kleinere. Das kleinste $\aleph$ ist $\aleph_0$ (die Mächtigkeit der abzählbaren Mengen).

Die Gesamtheit aller Ordinalzahlen, denen dieselbe Mächtigkeit $\aleph_\nu$ zukommt, nennt man eine *Zahlklasse* und bezeichnet sie mit $Z(\aleph_\nu)$. Unter allen Ordinalzahlen von $Z(\aleph_\nu)$ gibt es eine kleinste, sie heißt *die Anfangszahl von* $Z(\aleph_\nu)$ und wird mit Ω_ν bezeichnet. Alle Anfangszahlen, mit Ausnahme von ω, sind ε-Zahlen.

Die Mächtigkeiten der wohlgeordneten Mengen bilden, wenn man sie der Größe nach ordnet, selbst eine wohlgeordnete Menge; man kann daher die Alefs der Reihe nach bezeichnen mit:

$$\aleph_0, \aleph_1, \aleph_2, \ldots, \aleph_n, \ldots, \aleph_\omega, \aleph_{\omega+1}, \ldots, \text{etc.}$$

Die Gesamtheit aller Ordinalzahlen von $Z(\aleph_0)$ *bildet eine Menge von der Mächtigkeit* $\aleph_1$. Allgemein: *die Gesamtheit aller Ordinalzahlen, deren Mächtigkeit kleiner ist als* $\aleph_\nu$, *bildet eine Menge von der Mächtigkeit* $\aleph_\nu$.

Für die *Addition* und *Multiplikation der Alefs* gelten die Formeln:

Ist $\aleph_\mu \leqq \aleph_\nu$, so ist $\aleph_\mu + \aleph_\nu = \aleph_\nu$ und $\aleph_\mu \cdot \aleph_\nu = \aleph_\nu$.

Die hier in den §§ 5, 6, 7 skizzierte abstrakte Mengenlehre wurde begründet durch G. Cantor, auf den auch die meisten oben angeführten Theoreme zurückgehen (G. Cantor, *Grundlagen einer allgemeinen Mannigfaltigkeitslehre,* Leipzig 1883, *Über unendliche lineare Punktmannigfaltigkeiten, Math. Ann.* **15**, 1 (1879); **17**, 355 (1880), **20**, 113 (1882); **21**, 51, 545 (1883), **23**, 453 (1884), *Beiträge zur Begründung der transfiniten Mengenlehre, Math. Ann.* **46**, 481 (1895); **49**, 207 (1897), ein Teil von Cantors Abhandlungen findet sich in französischer Übersetzung in *Acta math.* **2**).

Neuere Darstellungen: A. Schoenflies, *Entwicklung der Lehre von den Punktmannigfaltigkeiten* (*Math.-Ver.* **8**, 1900) und G. Hessenberg, *Grundbegriffe der Mengenlehre* (Abh. der Friesschen Schule, neue Folge, Heft 4, Göttingen 1906). Ferner É. Borel, *Leçons sur la théorie des fonctions,* Paris 1900 und E. V. Huntington, *Ann. of math.* (2) **6**, 151 u. **7**, 15.

G. Cantor hat behauptet, daß jede beliebige Menge einer wohlgeordneten Menge äquivalent ist; ein von E. Zermelo veröffentlichter Beweis (*Math. Ann.* **59**, 514 (1904)) hat viele Einwendungen erfahren (hauptsächlich É. Borel, R. Baire, H. Lebesgue, *Bull. soc. math.* **33**, 261, A. Schoenflies, *Math. Ann.* **60**, 181, H. Poincaré, *Revue de métaphys. et de mor.* **14**, 294), auf die Zermelo (*Math. Ann.* **65**, 111 (1907)) eingehend geantwortet hat.

G. Cantor hat auch vermutet, daß die Mächtigkeit des Kontinuums gleich $\aleph_1$ sei, doch wurde ein bindender Beweis hierfür bisher nicht erbracht.

Aus der Betrachtung der Menge *aller* transfiniten Ordinalzahlen ergibt sich, wie zuerst C. Burali-Forti bemerkte (*Rend. Pal.* **11**, 154 (1897)), ein logischer Widerspruch. Über diesen, sowie über ähnliche Widersprüche vgl. B. Russell, *The principles of mathematics* **1**, 101 u. *Rev. de metaph. et de mor.* **14**, 627, A. Schoenflies, *Math.-Ver.* **15**, 19, G. Hessenberg, *l. c.,* 621, H. Poincaré, *l. c.,* G. Peano, *Rev. de math.* **8**, 149, Zermelo, *l. c.*

§ 8. Die Punktmengen.

Unter einer *linearen Punktmenge* versteht man eine Menge, deren Elemente Punkte einer Geraden sind; ebenso unter einer

ebenen (n-dimensionalen) Punktmenge eine Menge, deren Elemente Punkte einer Ebene (eines n-dimensionalen Raumes) sind. Wir werden im folgenden nur von linearen Punktmengen sprechen. Die meisten Sätze haben ihr Analogon auch für ebene und n-dimensionale Punktmengen.

Zufolge der Möglichkeit einer eineindeutigen Zuordnung zwischen den reellen Zahlen und den Punkten einer Geraden (§ 2) ist die Lehre von den linearen Punktmengen identisch mit der Lehre von den Mengen reeller Zahlen.

Unter dem *Intervall* (a, b) versteht man die Gesamtheit aller Punkte einer Geraden, die zwischen den Punkten a und b liegen. Ein Punkt des Intervalles (a, b) heißt innerer Punkt dieses Intervalles, wenn er von den Endpunkten a und b des Intervalles verschieden ist.

Ist jedem Punkte x des Intervalles (a, b), die Endpunkte inbegriffen, ein Intervall δ_x zugeordnet, in dessen Innern er liegt, so läßt sich aus der Menge der Intervalle δ_x eine endliche Anzahl von Intervallen so auswählen, daß sie das Intervall (a, b) vollständig überdecken (Satz von É. Borel, *Ann. éc. norm.* (3) **12**, 51 (1895); H. Lebesgue, *Lçç. sur l'intégration*, Paris (1904), 104).

Als *Komplementärmenge* einer linearen Punktmenge P bezüglich des Intervalles (a, b) bezeichnet man die Menge aller Punkte von (a, b), die nicht zu P gehören.

Ein Punkt O heißt *Häufungspunkt* (Verdichtungspunkt) der linearen Punktmenge P, wenn in jedem Intervalle, das den Punkt O als inneren Punkt enthält, sich ein von O verschiedener Punkt befindet, der der Menge P angehört. — Ein Häufungspunkt von P ist nicht notwendig selbst Punkt von P. Betrachtet man z. B. die Menge, die aus den Punkten mit den Abszissen

$$1, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{n}, \ldots$$

besteht, so ist der Nullpunkt Häufungspunkt dieser Menge, ohne ihr anzugehören.

Jede ganz in einem endlichen Intervall liegende unendliche Punktmenge besitzt in diesem Intervalle mindestens einen Häufungspunkt (Satz von B. Bolzano. Wegen dieser Bezeichnung vgl. G. Cantor, *Math. Ann.* **23**, 455 und O. Stolz, *Math. Ann.* **18**, 252).

Eine Punktmenge heißt *abgeschlossen*, wenn sie ihre sämtlichen Häufungspunkte enthält; sie heißt *isoliert*, wenn keiner ihrer Punkte ein Häufungspunkt ist; sie heißt *in sich dicht*,

wenn jeder ihrer Punkte ein Häufungspunkt ist; sie heißt *perfekt*, wenn sie abgeschlossen und in sich dicht ist. (Man beachte, daß diese Bezeichnungsweisen mit den in § 6 für Ordnungstypen eingeführten sich nicht decken.)

Eine perfekte Menge wird z. B. gebildet von jenen Punkten des Intervalles (0, 1), deren Abszissen sich als systematische Brüche von der Basis 3 (vgl. § 2) schreiben lassen, ohne Verwendung des Zählers 1. (Man beachte, daß zu diesen Abszissen etwa auch die Zahl $\frac{1}{3}$ gehört, da sie in der Form geschrieben werden kann: $\frac{2}{3^2} + \frac{2}{3^3} + \cdots + \frac{2}{3^n} + \cdots$)

Jede (unendliche) isolierte Menge ist abzählbar, eine abgeschlossene Menge ist entweder abzählbar oder von der Mächtigkeit des Kontinuums, jede perfekte Menge hat die Mächtigkeit des Kontinuums.

Jede abgeschlossene Menge läßt sich zerlegen in eine abzählbare und eine perfekte Menge (wobei auch einer dieser Bestandteile fehlen kann). (Cantor-Bendixsonsches Theorem; J. Bendixson, *Acta math.* 2, 415 (1883); der einfachste Beweis stammt von E. Lindelöf, *Acta math.* 29, 183.)

Die Komplementärmenge einer in (a, b) gelegenen abgeschlossenen Menge bezüglich dieses Intervalles besteht aus den inneren Punkten einer (endlichen oder abzählbaren) Menge sich nicht überdeckender Intervalle.

Eine Punktmenge heißt *überall dicht* im Intervalle (a, b), wenn sich in jedem beliebigen Teilintervalle von (a, b) Punkte der Menge finden. Sie heißt *nirgends dicht* in (a, b), wenn sie in keinem Teilintervall von (a, b) überall dicht ist.

Die Menge der Punkte mit rationalen Abszissen ist überall dicht, die oben angegebene perfekte Menge ist nirgends dicht im Intervalle (0, 1).

Die Punktmenge, die aus allen Häufungspunkten von P besteht, heißt die *erste Ableitung von* P und wird mit P' bezeichnet. *Sie ist stets abgeschlossen.*

Die Ableitung einer abgeschlossenen Menge ist in der Menge selbst enthalten; eine in sich dichte Menge ist in ihrer Ableitung enthalten; eine perfekte Menge ist mit ihrer Ableitung identisch.

Die Ableitung einer im Intervall (a, b) überall dichten Menge enthält sämtliche Punkte von (a, b).

Die erste Ableitung von P' wird zweite Ableitung von P genannt und mit P'' bezeichnet. Allgemein läßt sich, durch vollständige Induktion, die n^{te} Ableitung $P^{(n)}$ von P definieren

als die erste Ableitung der $(n-1)^{\text{ten}}$ Ableitung $P^{(n-1)}$. Alle Ableitungen $P^{(n)}$ von P sind, sofern sie überhaupt Punkte enthalten, abgeschlossene Mengen und $P^{(n)}$ ist in $P^{(n-1)}$ enthalten.

Enthalten alle Ableitungen $P^{(n)}$ von P Punkte, so gibt es Punkte, die allen $P^{(n)}$ gemeinsam sind, und zwar bilden diese Punkte eine abgeschlossene Menge, die mit $P^{(\omega)}$ bezeichnet und ω^{te} Ableitung von P genannt wird.

Ist α irgendeine Ordinalzahl und für jede kleinere Ordinalzahl β die β^{te} Ableitung $P^{(\beta)}$ von P definiert, so läßt sich auch eine α^{te} Ableitung $P^{(\alpha)}$ definieren, und zwar in folgender Weise: Ist α eine Grenzzahl, so ist $P^{(\alpha)}$ die Menge derjenigen Punkte, die allen $P^{(\beta)}$ gemeinsam sind; ist α keine Grenzzahl und β_0 die unmittelbar vorhergehende Zahl $(\alpha = \beta_0 + 1)$, so ist $P^{(\alpha)}$ die erste Ableitung von $P^{(\beta_0)}$. Da $P^{(\beta)}$ für $\beta \leqq \omega$ bereits definiert ist, so ist durch diese Festsetzungen $P^{(\alpha)}$ für alle Ordinalzahlen α definiert. *Jede so definierte Ableitung ist, wenn sie überhaupt Punkte enthält, eine abgeschlossene Menge, und wenn $\beta < \alpha$ ist, so ist $P^{(\alpha)}$ ganz in $P^{(\beta)}$ enthalten.*

Ist P eine beliebige Punktmenge, so gibt es stets eine endliche oder der Zahlklasse Z $(\aleph_0)$ angehörende Zahl α, so daß $P^{(\alpha)}$ entweder gar keinen Punkt enthält oder perfekt ist. Alle folgenden Ableitungen sind dann gleich $P^{(\alpha)}$. Enthält $P^{(\alpha)}$ gar keinen Punkt, so heißt P reduzibel.

Jede reduzible Menge ist abzählbar.

Der *Inhalt* einer Punktmenge wird in der folgenden Weise definiert. Sei P eine im Intervalle (a, b) liegende Punktmenge. Man betrachte eine endliche oder abzählbare Menge von Teilintervallen von (a, b), die die sämtlichen Punkte von P enthalten, und bilde die Summe S der Längen aller dieser Intervalle (die auch unendlich groß ausfallen kann). Macht man dies für alle möglichen endlichen und abzählbaren Intervallmengen der angegebenen Art, so hat die Menge der so gebildeten Zahlen S eine bestimmte Zahl zur unteren Grenze (vgl. § 10). Diese Zahl, die sicher nicht negativ und nicht größer als die Länge $b - a$ von (a, b) ist, wird der *äußere Inhalt* unserer Punktmenge P genannt. Nun bilde man den äußeren Inhalt der Komplementärmenge von P in bezug auf das Intervall (a, b) und subtrahiere ihn von der Länge $b - a$ dieses Intervalls. Die so erhaltene Zahl heißt der *innere Inhalt* von P. Er ist gewiß nicht größer als der äußere Inhalt. Sind äußerer und innerer Inhalt von P einander gleich, so heißt die Menge P *meßbar*, und der gemeinsame Wert von äußerem und innerem Inhalt wird kurz *der*

Inhalt von P genannt. (Bezüglich der Existenz nicht meßbarer Punktmengen vgl. H. Lebesgue, *Bull. soc. math.* **35**, 202.)

Ist die in (a, b) gelegene Punktmenge P meßbar, so ist auch ihre Komplementärmenge in bezug auf (a, b) meßbar, und die Summe der Inhalte dieser beiden Mengen ist $b - a$.

Daher: *Jede abgeschlossene in (a, b) gelegene Punktmenge P ist meßbar. Ihr Inhalt ist gleich der Länge von (a, b) vermindert um die Summe der Längen derjenigen Intervalle, welche von den nicht zu P gehörenden Punkten von (a, b) gebildet werden.*

Jede abzählbare Punktmenge ist meßbar und hat den Inhalt Null.

Sei eine endliche oder abzählbare Menge von meßbaren Punktmengen $P_1, P_2, \ldots, P_i, \ldots$ gegeben; *dann ist auch die Menge P, die aus sämtlichen Punkten von $P_1, P_2, \ldots, P_i, \ldots$ besteht, meßbar. Haben von den Mengen P_i keine zwei einen Punkt gemein, so ist der Inhalt von P gleich der Summe der Inhalte der P_i. Ebenfalls ist die Menge Q derjenigen Punkte, die allen P_i gemeinsam sind, meßbar, und ist für alle i die Menge P_i in P_{i-1} enthalten, so ist der Inhalt von Q gleich der Limite der Inhalte der P_i.*

Die Theorie der Punktmengen rührt in ihren Grundzügen ebenfalls von G. Cantor her. Außer der in § 7 erwähnten Literatur sei hier noch das Buch von N. H. und G. C. Young, *Theory of sets of points* (Cambridge 1906) genannt, wo man ein vollständiges Literaturverzeichnis findet. Die oben mitgeteilte Definition des Inhaltes einer Punktmenge rührt von H. Lebesgue her (*Ann. di mat.* (3) **7**, 231 (1902)); näheres darüber in dessen *Leçons sur l'intégration* (Paris 1904), 102 ff. Eine ähnliche Definition gibt W. H. Young, *Proc. Lond.* (2) **2**, 16 (1905) und *l. c.*, Chap. V. Weniger weittragende Definitionen des Inhaltes einer Punktmenge hatten vorher u. a. gegeben: G. Cantor, *Math. Ann.* **23**, 473 (1884); C. Jordan, *Journ. de math.* (4) **8**, 76 (1892); É. Borel, *Leçons sur la théorie des fonctions*, 46.

§ 9. Der Funktionsbegriff. Der Begriff des Grenzwertes.

Eine Größe y heißt eine (eindeutige) *Funktion* der reellen Veränderlichen x im Intervalle (a, b), wenn jedem in dieses Intervall fallenden Werte von x in eindeutiger Weise ein Wert von y zugeordnet ist.[1]) Allgemeiner heißt y eine Funktion von x

1) In diesem allgemeinen Sinne, der von der Möglichkeit einer analytischen Darstellung gänzlich absieht, erscheint der Funktions-

auf einer gegebenen linearen Punktmenge, wenn jedem einen Punkt dieser Menge darstellenden Werte von x ein Wert von y zugeordnet ist. Analog heißt y eine Funktion der n Veränderlichen $x_1, x_2, \ldots, x_n$ in einem Gebiete (oder auf einer Punktmenge) des n dimensionalen Raumes unserer n Veränderlichen, wenn jedem Wertsysteme der $x_1, x_2, \ldots, x_n$, das einen Punkt dieses Gebietes (bezw. dieser Punktmenge) darstellt, ein Wert von y zugeordnet ist.

Unter einem *Polynom in* x versteht man einen Ausdruck der Form:

$$a_0 x^k + a_1 x^{k-1} + \cdots + a_{k-1} x + a_k,$$

wo $a_0, a_1, \ldots, a_k$ Konstante bedeuten. Ist $a_0 \neq 0$, so heißt das Polynom *vom* k^{ten} *Grade*. Analog heißt die Summe einer endlichen Anzahl von Ausdrücken der Form:

$$a_{k_1, k_2, \ldots, k_n} x_1^{k_1} x_2^{k_2} \ldots x_n^{k_n}$$

ein *Polynom* in $x_1, x_2, \ldots, x_n$. Dabei bedeuten $k_1, k_2, \ldots, k_n$ nicht negative ganze Zahlen, $a_{k_1, k_2, \ldots, k_n}$ irgendwelche Konstante, und x_i^0 ist durch 1 zu ersetzen. Bildet man für alle Glieder, deren Koeffizient $a_{k_1, k_2, \ldots, k_n}$ nicht Null ist, den Ausdruck $k_1 + k_2 + \cdots + k_n$, so wird die größte der so erhaltenen Zahlen als der *Grad* unseres Polynomes bezeichnet.

Eine Funktion, die sich als Quotient zweier Polynome darstellen läßt, heißt eine *rationale Funktion*. Die Polynome heißen auch *ganze rationale* Funktionen.

Eine Funktion y von x heißt eine *algebraische Funktion* von x, wenn sie einer Gleichung der Form:

$$P_0(x) y^k + P_1(x) y^{k-1} + \cdots + P_{k-1}(x) y + P_k(x) = 0$$

genügt, wo die $P_i(x)$ Polynome in x bedeuten. Analog werden die algebraischen Funktionen von mehreren Veränderlichen definiert.

Eine Funktion $f(x_1, x_2, \ldots, x_n)$ der n Veränderlichen $x_1, x_2, \ldots, x_n$ heißt *homogen vom Grade* (der Ordnung) r, wenn für alle Werte von t die Relation besteht:

$$f(tx_1, tx_2, \ldots, tx_n) = t^r f(x_1, x_2, \ldots, x_n).$$

---

begriff zum erstenmal bei G. Lejeune Dirichlet (*Über die Entwicklung ganz willkürlicher Funktionen etc.* (1837), *Werke* **1**, 135). Frühere Mathematiker hatten das Wort *Funktion* in viel beschränkterem Sinne gebraucht. Vgl. *Differentialrechnung* § 2.

Eine Funktion $f(x)$ einer Veränderlichen heißt *monoton wachsend* im Intervall (a, b), wenn für irgend zwei Werte x' und x'' dieses Intervalles aus $x' > x''$ folgt: $f(x') \geqq f(x'')$. Sie heißt *monoton abnehmend*, wenn aus $x' > x''$ folgt: $f(x') \leqq f(x'')$. Die monoton wachsenden und monoton abnehmenden Funktionen werden zusammengefaßt unter dem Namen *monotone Funktionen*.

Man definiert: Die Funktion $f(x)$ hat für $x = a$ den *rechtsseitigen Grenzwert* A, in Zeichen: $\lim\limits_{x=a+0} f(x) = A$, wenn zu jeder beliebigen positiven Zahl ε eine positive Zahl η gehört, derart daß für alle der Ungleichung $0 < x - a < \eta$ genügenden x die Ungleichung besteht: $|f(x) - A| < \varepsilon$. Die Definition des *linksseitigen Grenzwertes* $\lim\limits_{x=a-0} f(x)$ erhält man, indem man die obige Ungleichung ersetzt durch: $0 > x - a > -\eta$.

Hat $f(x)$ für $x = a$ sowohl den rechtsseitigen als den linksseitigen Grenzwert A, so schreibt man: $\lim\limits_{x=a} f(x) = A$ und sagt: *$f(x)$ hat für $x = a$ den Grenzwert A.*

Man definiert weiter: $\lim\limits_{x=a+0} f(x) = +\infty$ (bezw. $= -\infty$), wenn zu jeder beliebigen Zahl B eine positive Zahl η gehört, derart daß für alle der Ungleichung $0 < x - a < \eta$ genügenden x die Ungleichung besteht: $f(x) > B$ (bezw. $< B$). In ganz analoger Weise wird definiert: $\lim\limits_{x=a-0} f(x) = +\infty$ und $\lim\limits_{x=a-0} f(x) = -\infty$. Ist gleichzeitig $\lim\limits_{x=a+0} f(x) = +\infty$ und $\lim\limits_{x=a-0} f(x) = +\infty$, so schreibt man wieder $\lim\limits_{x=a} f(x) = +\infty$. Analoge Bedeutung hat: $\lim\limits_{x=a} f(x) = -\infty$.

Ferner: $\lim\limits_{x=+\infty} f(x) = A$ (bezw. $\lim\limits_{x=-\infty} f(x) = A$), wenn zu jeder beliebigen positiven Zahl ε eine Zahl b gehört, derart daß für alle der Ungleichung $x > b$ (bezw. $x < b$) genügenden x die Ungleichung besteht: $|f(x) - A| < \varepsilon$. Analog: $\lim\limits_{x=+\infty} f(x) = +\infty$ (bezw. $= -\infty$), wenn zu jeder beliebigen Zahl B eine Zahl b gehört, derart daß für alle der Ungleichung $x > b$ genügenden x die Ungleichung $f(x) > B$ (bezw. $f(x) < B$) besteht. Und in ähnlicher Weise werden die Zeichen $\lim\limits_{x=-\infty} f(x) = +\infty$ (bezw. $= -\infty$) definiert.

Eine Beziehung zwischen dem hier eingeführten Grenzbegriff und dem in § 2 eingeführten Zeichen $\lim\limits_{n=\infty} a_n$ wird hergestellt durch die Sätze:

Aus $\lim\limits_{x=+\infty} f(x) = A$ folgt $\lim\limits_{n=\infty} f(n) = A$; aus $\lim\limits_{x=a} f(x) = A$

und $\lim_{n=\infty} x_n = a$ folgt: $\lim_{n=\infty} f(x_n) = A$, vorausgesetzt, daß (wenigstens von einem bestimmten Wert N des Index n an) alle x_n von a verschieden sind. Die Umkehrung dieser beiden Sätze gilt nicht.

Ist $\lim_{x=a+0} f_1(x) = A$ *und* $\lim_{x=a+0} f_2(x) = A$ *und ist eine dritte Funktion* $f(x)$ *so beschaffen, daß für alle der Ungleichung* $0 < x - a < \eta$ *genügenden* x (oder kurz gesprochen: in einer rechtsseitigen Umgebung von a) *die Funktion* $f(x)$ *der Größe nach zwischen* $f_1(x)$ *und* $f_2(x)$ *liegt, so ist auch* $\lim_{x=a+0} f(x) = A$.

Ist $\lim_{x=a+0} f_1(x) = +\infty$ (bezw. $= -\infty$) *und besteht in einer rechtsseitigen Umgebung von a die Ungleichung:* $f(x) \geqq f_1(x)$ (bezw. $f(x) \leqq f_1(x)$), *so ist auch* $\lim_{x=a+0} f(x) = +\infty$ (bezw. $= -\infty$).

Analoges gilt für $\lim_{x=a-0}$ und $\lim_{x=a}$.

Ist $f(x)$ *in einer rechtsseitigen Umgebung von a monoton wachsend (monoton abnehmend), so existiert stets* $\lim_{x=a+0} f(x)$; *bleibt außerdem* $f(x)$ *in dieser Umgebung von a größer (bezw. kleiner) als eine feste Zahl B, so ist* $\lim_{x=a+0} f(x)$ *eine endliche Zahl.*

Analoges gilt für $\lim_{x=a-0}$, $\lim_{x=+\infty}$ und $\lim_{x=-\infty}$. Speziell: *Ist eine Folge reeller Zahlen* a_n *gegeben, die sämtlich unter einer Zahl A liegen, und folgt* (wenigstens von einem bestimmten Werte des Index n an) *aus* $n' > n''$ *die Ungleichung:* $a_{n'} > a_{n''}$, *so existiert eine endliche Zahl a, derart daß* $\lim_{n=\infty} a_n = a$.

Ist $\lim_{x=a} f_1(x) = A_1$ und $\lim_{x=a} f_2(x) = A_2$, so ist

$$\lim_{x=a} (f_1(x) + f_2(x)) = A_1 + A_2; \quad \lim_{x=a} (f_1(x) - f_2(x)) = A_1 - A_2;$$

$$\lim_{x=a} f_1(x) \cdot f_2(x) = A_1 \cdot A_2$$

und, vorausgesetzt daß A_2 nicht Null ist:

$$\lim_{x=a} \frac{f_1(x)}{f_2(x)} = \frac{A_1}{A_2}.$$

Man sagt: Die Funktion $f(x, y)$ hat für $x = a$, $y = b$ den Grenzwert A, in Zeichen: $\lim_{x=a,\, y=b} f(x, y) = A$, wenn zu jeder positiven Zahl ε eine positive Zahl η gehört, derart daß für jedes von a, b verschiedene Wertepaar x, y, das den Ungleichungen genügt: $|x - a| < \eta$, $|y - b| < \eta$, die Ungleichung

besteht: $|f(x, y) - A| < \varepsilon$. Aus $\lim_{x=a,\, y=b} f(x, y) = A$ folgt $\lim_{x=a}(\lim_{y=b} f(x, y)) = A$ und $\lim_{y=b}(\lim_{x=a} f(x, y)) = A$ (vorausgesetzt, daß die auftretenden Limiten existieren), aber nicht umgekehrt.

Grenzwerte von *Doppelfolgen* $a_{m,n}$ (wo m und n jedes für sich die Reihe der natürlichen Zahlen durchlaufen) werden definiert durch: $\lim_{m,n=\infty} a_{m,n} = A$, wenn zu jeder positiven Zahl ε ein Wert M von m und ein Wert N von n gehört, so daß aus den beiden Ungleichungen $m > M$, $n > N$ folgt: $|a_{m,n} - A| < \varepsilon$. Näheres über Doppelfolgen bei A. Pringsheim, *Münch. Ber.* **27**, 103, *Math. Ann.* **53**, 289, F. London, *Math. Ann.* **53**, 322.

§ 10. Obere und untere Grenze. Stetigkeit und Unstetigkeit.

Sei irgendeine Menge von Zahlen Z gegeben. Wir teilen die rationalen Zahlen in zwei Klassen (vgl. § 2) nach der folgenden Vorschrift: in die erste dieser Klassen geben wir eine rationale Zahl r, wenn sämtliche Zahlen von Z kleiner als r sind; in die zweite dieser Klassen alle übrigen rationalen Zahlen. Enthält die erste dieser Klassen keine Zahl, so sagt man, *die Menge Z hat die obere Grenze* $+\infty$; enthalten beide Klassen Zahlen, so ist dadurch ein Schnitt im Gebiet der rationalen Zahlen definiert, und die diesen Schnitt erzeugende reelle Zahl M heißt die *obere Grenze der Menge Z*. Analog wird die untere Grenze m der Menge Z definiert, die, wenn sie nicht eine endliche Zahl ist, den Wert $-\infty$ hat. Der erste, der die Begriffe der oberen und unteren Grenze verwendete, war B. Bolzano (*Rein analytischer Beweis des Lehrsatzes etc.*, Prag 1817, herausg. in *Ostwalds Klassikern* No. 153 von Ph. E. B. Jourdain). Eine größere Verbreitung erhielten sie erst unter dem Einflusse von K. Weierstraß.

Sei eine Funktion $f(x)$ gegeben, die für alle Punkte des Intervalls (a, b) definiert ist, und Z sei die Menge der Werte, die $f(x)$ in diesem Intervalle annimmt. Die obere und untere Grenze von Z werden dann als *obere (untere) Grenze von* $f(x)$ *im Intervalle* (a, b) bezeichnet. Die Differenz zwischen oberer und unterer Grenze heißt *die Schwankung von* $f(x)$ *im Intervalle* (a, b).

Sei eine Folge von Intervallen δ_i $(i = 1, 2, \ldots, n, \ldots)$ gegeben, die alle den Punkt x_0 als inneren Punkt enthalten, und deren Länge mit wachsendem i gegen Null konvergiert,

und sei M_{δ_i} die obere Grenze von $f(x)$ in δ_i. Haben dann alle M_{δ_i} den Wert $+\infty$, so sagt man: *die Funktion $f(x)$ hat im Punkte x_0 die obere Grenze $+\infty$.* Im entgegengesetzten Falle existiert stets $\lim_{i=\infty} M_{\delta_i}$ und ist eine endliche Zahl; *sie heißt die obere Grenze von $f(x)$ im Punkte x_0.* Die so definierte obere Grenze ist unabhängig von der Wahl der Intervalle δ_i. Analog wird die untere Grenze von $f(x)$ im Punkte x_0 definiert. Die Differenz zwischen oberer und unterer Grenze im Punkte x_0 heißt *die Schwankung von $f(x)$ im Punkte x_0.*

Ist M die obere Grenze von $f(x)$ im Intervalle (a, b), so gibt es in diesem Intervalle mindestens einen Punkt x_0, in dem die obere Grenze von $f(x)$ gleich M ist. (Analog für die untere Grenze.)

Analoges gilt für Funktionen von n Veränderlichen, wenn man statt der Intervalle abgeschlossene n dimensionale Gebiete betrachtet.

Die Funktion $f(x)$ heißt *stetig im Punkte x_0*, wenn zu jeder positiven Zahl ε eine positive Zahl η gehört, so daß aus $|x - x_0| < \eta$ die Ungleichung folgt: $|f(x) - f(x_0)| < \varepsilon$.

Ist $\lim_{x=x_0+0} f(x) = f(x_0)$ und $\lim_{x=x_0-0} f(x) = f(x_0)$, so ist $f(x)$ stetig im Punkte x_0 und umgekehrt.

Ist nur $\lim_{x=x_0+0} f(x) = f(x_0)$ (bezw. $\lim_{x=x_0-0} f(x) = f(x_0)$), so heißt $f(x)$ *rechtsseitig* (bezw. *linksseitig*) *stetig* im Punkte x_0.

Damit $f(x)$ stetig sei im Punkte x_0 ist notwendig und hinreichend, daß die Schwankung von $f(x)$ im Punkte x_0 gleich Null sei.

Sind $f(x)$ und $\varphi(x)$ stetig im Punkte x_0, so ist auch $f(x) + \varphi(x)$, $f(x) - \varphi(x)$, $f(x) \cdot \varphi(x)$, und, vorausgesetzt daß $\varphi(x_0) \neq 0$ ist, $\frac{f(x)}{\varphi(x)}$ stetig im Punkte x_0.

Setzt man $\varphi(x_0) = y_0$ und ist die Funktion $\varphi(x)$ stetig im Punkte x_0, $f(y)$ stetig im Punkte y_0, so ist die Funktion $f(\varphi(x))$ stetig im Punkte x_0.

Die Funktion $f(x)$ heißt *stetig im Intervalle (a, b)*, wenn sie stetig ist in jedem inneren Punkte dieses Intervalles, rechtsseitig stetig im Punkte a und linksseitig stetig im Punkte b.

Ist die Funktion $f(x)$ stetig in (a, b), so ist sie für alle Punkte dieses Intervalles gegeben, wenn ihre Werte an einer im Intervalle (a, b) überall dicht liegenden Punktmenge, z. B. der Menge der rationalen Punkte von (a, b) gegeben sind.

Hieraus folgt, daß die Menge aller in (a, b) stetigen Funktionen die Mächtigkeit des Kontinuums hat, während die Mächtigkeit aller Funktionen gleich c^c (und mithin größer als c) ist.

Ist die Funktion $f(x)$ stetig in (a, b), so ist ihre obere (untere) Grenze in diesem Intervalle eine endliche Zahl, und es gibt in diesem Intervalle mindestens einen Punkt, in dem der Wert von $f(x)$ gleich ist dieser oberen (unteren) Grenze.

Ist die Funktion $f(x)$ stetig in (a, b) und ist E ein Wert, der zwischen $f(a)$ und $f(b)$ liegt, so gibt es in (a, b) einen Punkt, in dem $f(x)$ den Wert E annimmt. Kürzer gesprochen: Eine stetige Funktion kann nicht von einem Werte zu einem anderen übergehen, ohne alle dazwischen liegenden Werte anzunehmen. Die Umkehrung dieses Satzes gilt nicht (vgl. G. Darboux, *Ann. éc. norm.* (2) **4**, 109 (1875) und H. Lebesgue, *Leç. sur l'intégration*, 89).

Ist die Funktion $f(x)$ stetig in (a, b), so gehört zu jeder positiven Zahl ε eine positive Zahl η, derart, daß für irgend zwei Punkte x' und x'' von (a, b), für die $|x' - x''| < \eta$ ist, die Ungleichung besteht: $|f(x') - f(x'')| < \varepsilon$ (Satz von der *gleichmäßigen* Stetigkeit; der erste Beweis wurde publiziert von E. Heine, *Journ. f. Math.* **74**, 188 (1872)).

Jede im Intervalle (a, b) stetige Funktion läßt sich in eine in diesem Intervalle gleichmäßig konvergente (vgl. den Abschnitt über *Funktionentheorie*) *Reihe von Polynomen entwickeln* (Satz von K. Weierstraß, *Berl. Sitzungsber.* (1885), 633, 789; Näheres über die Entwicklung stetiger Funktionen in Reihen von Polynomen bei É. Borel, *Leçons sur les fonct. de variables réelles*, Paris 1905, Chap. IV).

Sei die Funktion $f(x)$ stetig im Intervall (a, b) und aus $x' > x''$ folge $f(x') > f(x'')$ (oder $f(x') < f(x'')$). Setzt man $f(a) = A$ und $f(b) = B$ und ist y eine beliebige Zahl des Intervalles (A, B), so gibt es in (a, b) eine und nur eine Zahl x, für die $f(x) = y$ wird. Ordnet man der Zahl y diese Zahl x zu, so ist dadurch im Intervalle (A, B) der Veränderlichen y eine Funktion $\varphi(y)$ definiert, welche *die zur Funktion $f(x)$ inverse Funktion heißt.* Sie ist stetig in (A, B) und aus $y' > y''$ folgt $\varphi(y') > \varphi(y'')$ (bezw. $\varphi(y') < \varphi(y'')$). Sie genügt der Relation: $\varphi(y) = \varphi(f(x)) = x$, woraus $f(\varphi(y)) = y$, d. h. die zur inversen Funktion inverse ist die ursprüngliche Funktion.

Jede Stelle x_0, an der $f(x)$ nicht stetig ist, heißt eine *Unstetigkeitsstelle* von $f(x)$, der Wert der Schwankung von $f(x)$

an der Stelle x_0 ist dann nicht Null; er wird auch als der *Unstetigkeitsgrad* von $f(x)$ an der Stelle x_0 bezeichnet.

Ist in jedem Punkte des Intervalles (a, b), die Endpunkte inbegriffen, der Unstetigkeitsgrad von $f(x)$ kleiner als A, so gehört zu jeder positiven Zahl ε eine positive Zahl η, derart daß für irgend zwei Punkte x' und x'' von (a, b), für die $|x' - x''| < \eta$ ist, die Ungleichung besteht: $|f(x') - f(x'')| < A + \varepsilon$.

Eine Unstetigkeitsstelle x_0 von $f(x)$ heißt *von erster Art*, wenn sowohl $\lim_{x=x_0+0} f(x)$ als $\lim_{x=x_0-0} f(x)$ existiert, sonst *von zweiter Art*. Sind speziell $\lim_{x=x_0+0} f(x)$ und $\lim_{x=x_0-0} f(x)$ gleich derselben endlichen Zahl, aber nicht gleich $f(x_0)$, so heißt die Unstetigkeit *hebbar*, weil sich dann $f(x)$ durch bloße Abänderung des Funktionswertes an der Stelle x_0 in eine im Punkte x_0 stetige Funktion verwandeln läßt.

Eine Funktion heißt *punktweise unstetig* im Intervalle (a, b), wenn ihre Stetigkeitsstellen in diesem Intervalle überall dicht liegen, sie heißt *total unstetig* im Intervalle (a, b), wenn sie in jedem Punkte von (a, b) unstetig ist. (Über die allgemeinste Verteilung der Stetigkeitsstellen einer Funktion $f(x)$ vgl. W. H. Young, *Wien. Ber.* **112**, Abt. 2ª (1903), 1307.)

Eine unstetige Funktion $f(x)$ heißt *von erster Klasse* im Intervalle (a, b), wenn sie in diesem Intervalle Limite stetiger Funktionen ist, d. h. wenn eine Gleichung $f(x) = \lim_{n=\infty} f_n(x)$ für jeden Punkt von (a, b) besteht, wobei mit $f_n(x)$ Funktionen bezeichnet sind, die in (a, b) stetig sind. Eine unstetige Funktion heißt *von zweiter Klasse*, wenn sie Limite von Funktionen erster Klasse ist, ohne selbst von erster Klasse zu sein. Allgemein heißt eine unstetige Funktion *von n^{ter} Klasse*, wenn sie Limite von Funktionen $(n-1)^{ter}$ Klasse ist, ohne selbst von $(n-1)^{ter}$ oder niedrigerer Klasse zu sein. Eine unstetige Funktion heißt *von ω^{ter} Klasse*, wenn sie Limite von Funktionen aus Klassen von endlicher Ordnung ist, ohne selbst einer dieser Klassen anzugehören. Ist π irgendeine transfinite Ordinalzahl der Zahlklasse $Z(\aleph_0)$, so heißt eine unstetige Funktion *von π^{ter} Klasse*, wenn sie Limite von Funktionen von niedrigerer als der π^{ten} Klasse ist, ohne selbst einer dieser Klassen anzugehören.

Es lassen sich Funktionen definieren, die einer beliebig gegebenen Klasse (deren Ordnung eine endliche Zahl oder eine Zahl der Zahlklasse $Z(\aleph_0)$ ist) *angehören; es lassen sich aber auch Funktionen definieren, die keiner dieser Klassen angehören.*

Die genannte Klassifizierung der Funktionen rührt her von R. Baire, *Ann. di mat.* (3) **3**, 68 (1899). Näheres hierüber: R. Baire, *Leçons sur les fonctions discontinues,* Paris 1905, wo man die notwendige und hinreichende Bedingung dafür findet, daß eine Funktion der ersten Klasse angehöre, und *Acta math.* **30**, 1. Ferner H. Lebesgue, *Journ. de math.* (6) **1**, 139.

Ähnliche Definitionen und Sätze, wie die hier für Funktionen einer Veränderlichen angeführten, gelten für Funktionen von mehreren Veränderlichen, z. B. heißt eine Funktion von zwei Veränderlichen x, y stetig an der Stelle x_0, y_0, wenn zu jeder positiven Zahl ε eine positive Zahl η gehört, so daß aus dem gleichzeitigen Bestehen der beiden Ungleichungen $|x-x_0|<\eta$ und $|y-y_0|<\eta$ folgt: $|f(x, y)-f(x_0, y_0)|<\varepsilon$; oder, was dasselbe ist, wenn $\lim_{x=x_0,\, y=y_0} f(x, y) = f(x_0, y_0)$ ist.

Aus der Stetigkeit von $f(x, y_0)$ als Funktion von x an der Stelle x_0 und der Stetigkeit von $f(x_0, y)$ als Funktion von y an der Stelle y_0 folgt nicht die Stetigkeit von $f(x, y)$ an der Stelle x_0, y_0, wie man z. B. am Verhalten der Funktion $\frac{xy}{x^2+y^2}$ im Punkte $x=0$, $y=0$ erkennen kann.

Ist die Funktion $f(x_1, x_2, \ldots, x_n)$ in allen Punkten eines Gebietes der n Veränderlichen $x_1, x_2, \ldots, x_n$ stetig nach jeder einzelnen dieser Veränderlichen, so ist sie in diesem Gebiete höchstens von $(n-1)^{ter}$ Klasse (H. Lebesgue, *Bull. sciences math.* (2) $\mathbf{22}_1$, 284 (1898) und *Journ. de math.* (6) **1**, 201; siehe auch R. Baire, *Ann. di mat.* (3) **3**, 87).

Näheres über Stetigkeit und Unstetigkeit findet man, außer in den bereits genannten Werken, bei A. Schoenflies, *Entwicklung der Lehre von den Punktmannigfaltigkeiten* (*Math.-Ver.* **8**) und A. Pringsheim, *Enzykl.* **II** A 1, wo sich auch viele Literaturnachweise finden. Ferner: U. Dini, *Grundlagen für eine Theorie der Funktionen einer veränderlichen reellen Größe,* deutsch von J. Lüroth u. A. Schepp, Leipzig 1892; O. Stolz u. J. A. Gmeiner, *Einleitung in die Funktionentheorie,* Leipzig 1904 (*Teubners Sammlung* **14**), J. Pierpont, *The theory of functions of real variables,* Boston 1905. Zahlreiche Beispiele bei E. Pascal, *Esercizi e note critiche etc.,* Mailand 1895.

Sei die Funktion $f(x)$ definiert in (a, b), und sei eine endliche Anzahl beliebiger Punkte von (a, b) gegeben:

$$a < x_1 < x_2 < \cdots < x_n < b.$$

Wir bilden:

$$V = |f(x_1) - f(a)| + |f(x_2) - f(x_1)| + \cdots + |f(b) - f(x_n)|.$$

Die obere Grenze aller so erhältlichen Zahlen V wird bezeichnet als *totale Schwankung von* $f(x)$ *im Intervalle* (a, b). Ist diese obere Grenze endlich, so heißt $f(x)$ *von beschränkter Schwankung* (*à variation bornée*) *im Intervall* (a, b).

Jede in (a, b) *monotone Funktion ist daselbst von beschränkter Schwankung; und zwar ist ihre totale Schwankung* $|f(b) - f(a)|$. *Eine Funktion, die in* (a, b) *von beschränkter Schwankung ist, läßt sich darstellen als Summe zweier in* (a, b) *monotoner Funktionen. Ist eine Funktion in* (a, b) *von beschränkter Schwankung, und besitzt sie in diesem Intervall Unstetigkeitsstellen, so sind dieselben alle von erster Art und bilden eine endliche oder abzählbare Menge.*

Die Funktionen von beschränkter Schwankung wurden eingeführt von C. Jordan (*Cours d'analyse*, 2. éd., **1**, 54). Siehe auch E. Study, *Math. Ann.* **47**, 298 und H. Lebesgue, *Leç. sur l'intégration*, Chap. IV.

§ 11. Potenzen. Logarithmen. Wichtige Grenzwerte.

Bedeutet n eine natürliche Zahl, a eine beliebige, von Null verschiedene reelle Zahl, so versteht man unter a^n (der n^{ten} Potenz von a) das Produkt von n einander gleichen Faktoren a, unter a^{-n} die Zahl $\frac{1}{a^n}$, unter a^0 die Zahl 1. Ist ferner a positiv, so gibt es eine und nur eine positive Zahl, die zur n^{ten} Potenz erhoben a ergibt; sie heißt die positive n^{te} Wurzel aus a und wird mit $a^{\frac{1}{n}}$ oder $\sqrt[n]{a}$ bezeichnet; ist dann $r = \frac{m}{n}$ eine positive rationale Zahl, so versteht man unter a^r die positive n^{te} Wurzel aus a^m, und unter a^{-r} die Zahl $\frac{1}{a^r}$. Sind r und r' irgendwelche rationale Zahlen, so ist:

$$(1) \qquad a^r \cdot a^{r'} = a^{r+r'}; \quad (a^r)^{r'} = a^{r \cdot r'}$$

und wenn $r > r'$ ist:

$$(2) \qquad \begin{aligned} a^r &> a^{r'} \quad \text{für} \quad a > 1; \\ a^r &< a^{r'} \quad \text{für} \quad a < 1. \end{aligned}$$

Sei a positiv und > 1 (bezw. < 1) und ϱ eine beliebige irrationale Zahl. Um a^ϱ zu definieren, teilen wir die rationalen

Zahlen in zwei Klassen; in die eine Klasse geben wir alle rationalen Zahlen, die kleiner sind als irgendeine Zahl a^r, wo r eine rationale Zahl bedeutet, die kleiner (bezw. größer) als ϱ ist; in die andere Klasse geben wir alle übrigen rationalen Zahlen. Dadurch ist ein Schnitt im Gebiete der rationalen Zahlen definiert; die ihn erzeugende reelle Zahl wird mit a^ϱ bezeichnet. Die Gleichungen (1) und die Ungleichungen (2) gelten dann für beliebige reelle r und r'.

Hierdurch ist die Funktion a^x $(a > 0)$ für alle reellen x definiert; sie wird *Exponentialfunktion* genannt; sie ist stetig und wächst von 0 bis $+\infty$ oder nimmt ab von $+\infty$ bis 0, wenn x von $-\infty$ bis $+\infty$ wächst, je nachdem a größer oder kleiner als 1 ist. Die inverse Funktion (vgl. § 10) ist daher definiert von 0 bis $+\infty$; sie wird bezeichnet mit $\overset{a}{\log} x$ *(Logarithmus von x für die Basis a)*; sie ist stetig und wächst von $-\infty$ bis $+\infty$ oder nimmt ab von $+\infty$ bis $-\infty$, wenn x von 0 nach $+\infty$ wächst, je nachdem die Basis a größer oder kleiner als 1 ist.[1])

Es gelten die Formeln:

$$\log(x_1 x_2) = \log x_1 + \log x_2; \quad \log x^\varrho = \varrho \log x.$$

Der Übergang von einer Basis a zu einer anderen Basis a' wird vermittelt durch die Formel:

$$\overset{a'}{\log} x = \frac{\overset{a}{\log} x}{\overset{a}{\log} a'}.$$

Als *natürliche Logarithmen* bezeichnet man die Logarithmen von der Basis e, wobei e definiert ist durch:

$$e = \lim_{n=\infty} \left(1 + \frac{1}{n}\right)^n.$$

Näheres über Exponentialfunktion und Logarithmen im Abschnitte über *Funktionentheorie.*

Es gelten die Formeln:

$$\lim_{n=\infty} \left(1 + \frac{x}{n}\right)^n = e^x; \qquad \lim_{n=\infty} \left\{n\left(x^{\frac{1}{n}} - 1\right)\right\} = \overset{e}{\log} x$$

$$\lim_{x=+\infty} x^{\frac{1}{x}} = 1; \qquad \lim_{x=0} \frac{\lg(1 + ax)}{x} = a.$$

---

1) Im Falle $a = 1$ setzt man $1^x = 1$ für alle reellen x. Eine inverse Funktion kann in diesem Falle nicht gebildet werden.

Für jedes reelle k ist:

$$\lim_{x=+\infty} e^x \cdot x^k = +\infty; \qquad \lim_{x=+\infty} e^{-x} \cdot x^k = 0.$$

Für jedes positive k ist:

$$\lim_{x=+\infty} x^{-k} \cdot \log x = 0.$$

Die allgemeinste stetige Funktion, die der Funktionalgleichung

$$f(x+y) = f(x) + f(y)$$

genügt, ist $f(x) = a \cdot x$, *wo* a *eine Konstante bedeutet* (über unstetige Lösungen dieser Funktionalgleichung siehe R. Volpi, *Giorn. mat.* **35**, 104 (1894), G. Hamel, *Math. Ann.* **60**, 459 (1905); vgl. auch H. Lebesgue, *Atti Torino* **42**, 532 und F. Bernstein, *Math. Ann.* **64**, 417).

Die allgemeinste stetige Funktion, die der Funktionalgleichung

$$f(x+y) = f(x) \cdot f(y)$$

genügt, ist $f(x) = A^{ax}$, *wo* A *und* a *Konstante bedeuten.*

Die allgemeinste stetige Funktion, die der Funktionalgleichung

$$f(xy) = f(x) + f(y)$$

genügt, ist $f(x) = A \overset{a}{\log} x$.

Die allgemeinste stetige Funktion, die der Funktionalgleichung:

$$f(xy) = f(x) \cdot f(y)$$

genügt, ist $f(x) = x^a$.

Die vier zuletzt angeführten Sätze stammen von A. Cauchy (*Analyse algébrique* (1823), *Œuvres* (2) **3**, 98).

Es seien noch einige häufig vorkommende Grenzwertformeln zusammengestellt:

$$\lim_{x=0} \frac{\sin x}{x} = 1 \qquad \lim_{x=0} x \cdot \sin \frac{1}{x} = 0$$

$$\lim_{x=0} \frac{1-\cos x}{x} = 0 \qquad \lim_{x=0} \frac{\operatorname{tg} x}{x} = 1$$

$$\lim_{n=\infty} \frac{h^n}{n!} = 0 \qquad \lim_{x=0} \frac{(1+x)^k - 1}{x} = k$$

$$\lim_{n=\infty} \frac{n!}{n^n e^{-n} \sqrt{2\pi n}} = 1 \text{ (Stirlingsche Formel)}$$

$$\lim_{n=\infty} \frac{1^r + 2^r + \cdots + n^r}{n^{r+1}} = \frac{1}{r+1}, \text{ wenn } r+1 > 0;$$

für nichtpositives $r+1$ ist dieser Grenzwert unendlich

$$\lim_{n=\infty} \frac{1+\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{n}}{n} = 0.$$

Seien a und b reelle positive Zahlen. Bildet man der Reihe nach

$$a_1 = \tfrac{1}{2}(a+b) \qquad b_1 = \sqrt[+]{a\,b}$$
$$a_2 = \tfrac{1}{2}(a_1+b_1) \qquad b_2 = \sqrt[+]{a_1 b_1}$$
$$a_3 = \tfrac{1}{2}(a_2+b_2) \qquad b_2 = \sqrt[+]{a_2 b_2}$$
$$\cdots\cdots\cdots \qquad \cdots\cdots$$

so ist $\lim_{n=\infty} a_n = \lim_{n=\infty} b_n$, *und dieser Grenzwert heißt das arithmetisch-geometrische Mittel aus a und b* (*K. F. Gauss, Werke* **3**, 361).

Ist $\lim_{n=\infty} a_n = a$ *und setzt man* $s_n = \frac{a_1+a_2+\cdots+a_n}{n}$, *so ist auch* $\lim_{n=\infty} s_n = a$. Näheres über diesen und ähnliche Sätze: E. Cesàro, *Lehrbuch der algebraischen Analysis* (Leipzig 1904), 96, und É. Borel, *Leçons sur les séries divergentes* (Paris 1901), 87.

Aus $\lim_{x=+\infty} [f(x+1) - f(x)] = A$ *folgt* $\lim_{x=+\infty} \frac{f(x)}{x} = A$, *vorausgesetzt, daß obere und untere Grenze von $f(x)$ in jedem endlichen Intervalle endliche Zahlen sind* (A. Cauchy, *Analyse algébrique* (1823), *Œuvres* (2) **3**, 54; vgl. Stolz-Gmeiner, *Einleitung in die Funktionentheorie*, 31).

Aus $\lim_{x=+\infty} \frac{f(x+1)}{f(x)} = A$ *folgt* $\lim_{x=+\infty} (f(x))^{\frac{1}{x}} = A$, *vorausgesetzt, daß obere und untere Grenze von $f(x)$ in jedem endlichen Intervalle endliche positive Zahlen sind* (Cauchy, *l. c.*, 58).

Kapitel II.

Kombinatorik, Determinanten und Matrices.

Von *Alfred Loewy* in Freiburg i. Br.

§ 1. Die Lehre von den Kombinationen. Die Binomialkoeffizienten. Die Polygonal- und Pyramidalzahlen. Die figurierten Zahlen. Die Polynomialkoeffizienten.

Die Anzahl der verschiedenen Arten, auf die man n verschiedene Gegenstände in alle möglichen Reihenfolgen bringen kann, heißt *die Anzahl der Permutationen der n Gegenstände* (das Wort ist eingeführt von Jacob Bernoulli in der *Ars conjectandi* (1713), deutsche Ausgabe von Haußner in *Oswalds Klassikern der exakten Wissenschaften* No. 107, S. 78). Sie beträgt:

$$\Pi(n) = 1 \cdot 2 \ldots (n-1)n.$$

Das Produkt aller ganzen Zahlen von 1 bis n bezeichnet man auch (Kramp, *Éléments d'arithmétique universelle,* Cologne 1808) mit

$$n! = 1 \cdot 2 \cdot 3 \cdots n$$

und liest *n-Fakultät.*

Über das Wachsen von $n!$ gibt folgender Satz Aufschluß: *Ist c irgendeine ganze positive Zahl* > 1, *so kann man stets eine derartige positive ganze Zahl n finden, daß für alle ganzzahligen* $N > n$ *der Ausdruck* $N! > c^N$ *ist.*

Hat man für die n Gegenstände eine gewisse Reihenfolge festgesetzt und ordnet sie anders an, so spricht man jedesmal, wenn in dieser Anordnung zwei Gegenstände in der umgekehrten Reihenfolge wie bei der ursprünglichen Anordnung aufeinander folgen, von einer *Inversion* oder einem *Dérangement* (Gabriel Cramer, *Introduction à l'analyse des lignes courbes,* 1750,

Append., S. 658). Sind die Gegenstände mit 1, 2, 3, ..., n bezeichnet, so hat man in einer Anordnung jedesmal dann eine Inversion, wenn eine höhere Zahl einer niedrigeren voraufgeht.

Vertauscht man in einer gegebenen Anordnung irgend zwei Elemente miteinander, so spricht man von einer *Transposition* (Cauchy, *J. éc. polyt.*, Cah. 17, 18 (1815), *Œuvres* (2) 1, 81).

Eine Anordnung heißt *gerade* oder *von der ersten Klasse*, wenn sie eine gerade Anzahl von Inversionen enthält; im anderen Fall heißt sie *ungerade* oder *von der zweiten Klasse*. Eine Anordnung der ersten (zweiten) Klasse läßt sich stets aus der ursprünglichen oder natürlichen Anordnung durch eine gerade (ungerade) Anzahl von Transpositionen gewinnen. Jede Transposition ändert die Klasse einer Anordnung. Zwei beliebige Anordnungen können stets nur durch eine gerade oder nur durch eine ungerade Anzahl von Transpositionen ineinander übergeführt werden. Jede Anordnung kann durch Transpositionen in jede andere übergeführt werden. Irgend zwei Anordnungen gleicher (ungleicher) Klasse gehen durch eine gerade (ungerade) Anzahl Transpositionen ineinander über.

Über die *Klasse einer Anordnung* gibt auch das aus den n Größen $a_1, a_2, \ldots, a_n$ gebildete *Differenzenprodukt* (vgl. § 4 dieses Kapitels bei der Vandermondeschen oder Cauchyschen Determinante, S. 68)

$$\Delta(a_1, a_2, \ldots, a_n) = \prod_{i,j} (a_j - a_i) \qquad (j > i)$$

aus $\frac{n(n-1)}{2}$ Faktoren Aufschluß. Ist $x_1, x_2, \ldots, x_n$ irgendeine Anordnung der ersten n Zahlen 1, 2, ..., n, so gehört die Anordnung $x_1, x_2, \ldots, x_n$ in die erste (zweite) Klasse, falls der Quotient

$$\frac{\Delta(x_1, x_2, \ldots, x_n)}{\Delta(1, 2, \ldots, n)}$$

den Wert $+1$ (-1) hat. Die durch die Anordnung auf obige Weise bestimmte Zahl ± 1 heißt der *Modul der Anordnung;* sie ist für die Determinantentheorie von größter Wichtigkeit. Der Modul der Anordnung $x_1, x_2, \ldots, x_n$ kann auch auf folgende Weise bestimmt werden: er hat den Wert $(-1)^{n-r}$, wenn die Permutation

$$\begin{pmatrix} 1 & 2 & 3 & \ldots & n \\ x_1 & x_2 & x_3 & \ldots & x_n \end{pmatrix}$$

als Produkt r zyklischer Faktoren, von denen keine zwei ein Element gemeinsam haben, dargestellt werden kann; hierbei

sind auch die eingliedrigen Zykeln mitzurechnen (vgl. Kap. III). Alle angeführten Sätze finden sich bereits bei Cauchy, *J. éc. polyt.*, Cah. **17**, 29 (1815), *Œuvres* (2) **1**, 91ff. Von neueren Darstellungen vgl. man etwa Kronecker, *Vorl. üb. d. Theorie d. Determinanten,* herausg. von Hensel, Leipzig 1903, S. 299.

Es gibt $\frac{n!}{2}$ *Anordnungen der ersten Klasse und ebensoviele der zweiten Klasse.*

Die Anzahl der verschiedenen Anordnungen von n Elementen, unter denen a gleiche einer Art, b gleiche einer zweiten Art usw. vorkommen, ist $\frac{n!}{a!\,b!\,c!\,\ldots}$.

Die Anzahl der verschiedenen Arten, auf die man unter n gegebenen Gegenständen k auswählen kann, ohne die Reihenfolge zu berücksichtigen, in der die k Dinge gewählt werden, heißt die Anzahl *der einfachen Kombinationen der n Gegenstände zur* k^{ten} *Klasse.* Die Bezeichnung geht auf Blaise Pascal zurück, der von multitude des combinaisons de k dans n spricht (*Usage du triangle arithmétique,* Paris 1665, *Œuvres* **5**, 25 (1779), Cantor, *Vorl. über Geschichte der Mathematik,* 2. Aufl., Leipzig 1900, **2**, 752).

Die fragliche Anzahl beträgt

$$C_{n,k} = \frac{n(n-1)(n-2)\ldots(n-k+1)}{1\cdot 2\cdot 3\cdots k} = \binom{n}{k}\cdot$$

$$= \frac{n!}{k!\,(n-k)!}\cdot$$

Hieraus ergiebt sich unmittelbar der Satz:

$$C_{n,k} = C_{n,n-k}\cdot$$

Die Zahl $\binom{n}{k}$ läßt sich auch als die Anzahl der verschiedenen Mengen definieren, die man erhält, wenn man aus n verschiedenen Dingen Mengen von je k Dingen bildet. Da in keiner der $\binom{n}{k}$ Mengen ein Gegenstand mehrfach auftritt, so spricht man von *Kombinationen von n Dingen zu je k ohne Wiederholung.*

Wenn in den Kombinationen jedes Element wiederholt werden kann, so erhält man *die Kombinationen mit Wiederholung.* Die Anzahl der Mengen von je k Zahlen, die nur die Zahlen $1, 2, 3, \ldots, n$, die gleiche Zahl aber auch wiederholt, enthalten, oder, wie man auch sagt, die *Anzahl der Kombinationen von n Dingen zu je k mit Wiederholung, ist:*

$$C'_{n,k} = \frac{(n+k-1)(n+k-2)\dots(n+1)\,n}{1\cdot 2\cdot \dots\cdot (k-1)\,k} =$$

$$= \frac{(n+k-1)!}{k!\,(n-1)!} = C_{n+k-1,\,k} = \binom{n+k-1}{k}.$$

Vertauscht man die Elemente der einzelnen Kombinationen auf alle möglichen Weisen, so erhält man die *Variationen*. Die Anzahl der verschiedenen Arten, auf die man k Gegenstände, die man beliebig unter n gegebenen Gegenständen auswählt ($k \leqq n$), auf k feste Plätze verteilen kann, heißt die Anzahl *der einfachen Variationen der n Gegenstände zur k^{ten} Klasse*. Ihre Anzahl beträgt

$$D_{n,k} = n\,(n-1)\cdots(n-k+1) = \frac{n!}{(n-k)!} = \binom{n}{k} k!.$$

Für $k = n$ werden die *Variationen zu Permutationen*. Wenn in den Variationen jedes Element wiederholt werden darf, so erhält man *die Variationen mit Wiederholung*. Ihre Anzahl ist

$$D'_{n,k} = n^k.$$

Die Zahlen $C_{n,k} = \binom{n}{k}$ nennt man auch die *Binomialkoeffizienten*, weil sie die numerischen Koeffizienten der verschiedenen Glieder in der Entwicklung der Potenz eines Binoms sind:

$$(a+b)^n = a^n + \binom{n}{1} a^{n-1} b + \binom{n}{2} a^{n-2} b^2 +$$

$$\cdots + \binom{n}{k} a^{n-k} b^k \cdots + b^n.$$

Sämtliche Binomialkoeffizienten lassen sich mittels des sogenannten *Blaise Pascalschen arithmetischen Dreiecks* (*Traité du triangle arithmétique*, Paris 1665, *Œuvres* 5, 1 (1779))

1
1 1
1 2 1
1 3 3 1
1 4 6 4 1
1 5 10 10 5 1
. . . . . .

bilden. Man erhält die Zahlen einer Zeile durch Addition der beiden in der vorhergehenden Zeile unmittelbar über ihnen stehenden. Die Zahlen der $n + 1^{ten}$ Horizontalreihe geben die Binomialkoeffizienten der Entwicklung von $(a+b)^n$. Das Pascal-

sche Dreieck findet sich übrigens schon in einer chinesischen Schrift des 14. Jahrhunderts (Tropfke, *Geschichte der Elementar-Mathematik* 2, 326, Leipzig 1903).

Bei dem Symbol $\binom{n}{k}$ nennen wir n die *Basis* und k den *Index*. Das Symbol $\binom{n}{k}$, das bis jetzt nur für ganzzahlige positive n und k verwandt wurde, behält auch für *beliebige reelle und komplexe* n einen Sinn; $\binom{n}{0}$ sei als 1 definiert. Es ist

$$\binom{-n}{k} = (-1)^k \cdot \binom{n+k-1}{k}$$

und im besonderen

$$\binom{-1}{k} = (-1)^k.$$

Die Binomialkoeffizienten mit ganzen negativen Zahlen als Basis sind, abgesehen vom Vorzeichen, die in einer Diagonale stehenden Zahlen des Pascalschen arithmetischen Dreiecks.

Für beliebige m und n gilt das *Additionstheorem der Binomialkoeffizienten:*

$$\binom{n+m}{k} = \binom{n}{k}\binom{m}{0} + \binom{n}{k-1}\binom{m}{1} + \binom{n}{k-2}\binom{m}{2} + \cdots + \binom{n}{0}\binom{m}{k}.$$

(Euler, *Acta Acad. Petrop.* V (1781), S. 74.) Für $m = 1$ erhält man die grundlegende Eigenschaft des Pascalschen Dreiecks:

$$\binom{n+1}{k} = \binom{n}{k} + \binom{n}{k-1}.$$

(Michael Stifel, *Arithmetica integra,* Nürnberg 1544, zu vgl. M. Cantor, *Vorl. über Geschichte d. Math.* 2, 433). Für $m = -1$ wird:

$$(-1)^k \cdot \binom{n-1}{k} = \binom{n}{0} - \binom{n}{1} + \binom{n}{2} - \binom{n}{3} \cdots + (-1)^k \cdot \binom{n}{k}.$$

Für ganzzahlige positive $r < k$ ist $\binom{r}{k} = 0$; daher ergibt sich für ganzzahlige positive $n = k$ aus der letzten Formel:

$$0 = \binom{n}{0} - \binom{n}{1} + \binom{n}{2} - \binom{n}{3} \cdots (-1)^n \cdot \binom{n}{n}.$$

Für ganzzahlige positive n sei noch angeführt

$$\binom{n}{n} = 1, \quad \binom{n}{k} = \binom{n}{n-k},$$

ferner die aus dem Additionstheorem der Binomialkoeffizienten für ganzzahlige positive $n = m = k$ sich ergebende Formel:

$$\binom{2n}{n} = 1 + \binom{n}{1}^2 + \binom{n}{2}^2 + \binom{n}{3}^2 + \cdots + \binom{n}{n}^2.$$

Aus

$$(a+b)^n = a^n + \binom{n}{1} a^{n-1}b + \binom{n}{2} a^{n-2}b^2 + \cdots + \binom{n}{n} b^n$$

folgt durch Spezialisierung für $a = b = 1$:

$$2^n = 1 + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{n}.$$

Für ganzzahlige n gelten ferner die Formeln:

$$\binom{n}{k} - \binom{n}{k+1}\binom{k+1}{k} + \binom{n}{k+2}\binom{k+2}{k} \cdots (-1)^{n-k}\binom{n}{n}\binom{n}{k} = 0,$$

$$1 - \binom{n}{1}^2 + \binom{n}{2}^2 - \binom{n}{3}^2 \cdots (-1)^n \binom{n}{n}^2$$

ist gleich $(-1)^{\frac{n}{2}}\binom{n}{\frac{n}{2}}$ für gerade n und 0 für ungerade n.

Eine Sammlung von Summenformeln für die Binomialkoeffizienten gibt J. G. Hagen in seiner *Synopsis der höheren Mathematik* **1**, 64, Berlin 1891. Man vgl. hierzu auch das letzte Kapitel in dem *Lehrbuch der Kombinatorik* von E. Netto (*Teubners Sammlung* **7**, Leipzig 1901), das eine sehr vollständige Behandlung aller Fragen der Kombinatorik liefert. Als ältere ausführliche Darstellung der Kombinationslehre sei der zweite Teil der schon oben zitierten *Ars conjectandi* von J. Bernoulli genannt. Vorgänger von J. Bernoulli, die sich mit Kombinatorik beschäftigt haben, sind Bl. Pascal (*Traité du triangle arithmétique,* 1665), Leibniz (*Dissertatio de arte combinatoria,* 1666, zu vgl. Cantor, *Vorles. über Geschichte d. Math.*, 2. Aufl., 1901, **3**, 43 u. folg.), Wallis (*Treatise of algebra,* 1685).

Für die Binomialkoeffizienten mit der Basis $-\frac{1}{2}$ gelten die bemerkenswerten Ausdrücke:

$$\binom{-\frac{1}{2}}{1} = -\frac{1}{2}, \qquad \binom{-\frac{1}{2}}{2} = \frac{1\cdot 3}{2\cdot 4},$$

$$\binom{-\frac{1}{2}}{3} = -\frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}, \quad \binom{-\frac{1}{2}}{4} = \frac{1\cdot 3\cdot 5\cdot 7}{2\cdot 4\cdot 6\cdot 8}, \text{ usw.}$$

Über aus Binomialkoeffizienten gebildete Determinanten vgl. dieses Kapitel, § 4, ferner Pascal, *die Determinanten*

(*Teubners Sammlung* **3**, Leipzig 1900, S. 132), Netto, *Lehrb. der Kombinatorik*, S. 256.

Eine ganze positive Zahl N läßt sich immer und nur auf eine einzige Art als Summe von n Binomialkoeffizienten ausdrücken, deren Indices gegeben sind und aus den natürlichen Zahlen von 1 bis n bestehen, während von ihren Basen diejenige die kleinere ist, die zu einem kleineren Index gehört, d. h. die Formel:

$$N = \binom{x_1}{1} + \binom{x_2}{2} + \cdots + \binom{x_n}{n} \quad (x_k < x_{k+1})$$

hat immer eine und nur eine Auflösung in ganzen positiven Zahlen $x_1, x_2, \ldots, x_n$.

Bezeichnet man mit dem Symbol

$$\left[\begin{matrix} n \\ k \end{matrix}\right]$$

die Zahl $\binom{n+k-1}{k}$, so erhält man den weiteren Satz:

Wenn eine ganze positive Zahl N gegeben ist, so hat die Formel:

$$N = \left[\begin{matrix} 2 \\ x_1 \end{matrix}\right] + \left[\begin{matrix} 3 \\ x_2 \end{matrix}\right] + \cdots + \left[\begin{matrix} n+1 \\ x_n \end{matrix}\right],$$

in der

$$x_k < x_{k+1}$$

ist, immer nur eine Auflösung in ganzen positiven Zahlen $x_1, x_2, \ldots, x_n$. (Vgl. Pascal, *Giorn. di mat.* **25**, 48 (1887).)

Verallgemeinert man die Konstruktion des Pascalschen arithmetischen Dreiecks, so gelangt man zu den *Polygonal-* und *Pyramidalzahlen.*

Man bilde das Dreieck:

$$\begin{array}{ccccccccc}
 & & & & \delta & & 1 & & \\
 & & & \delta & & 1+\delta & & 1 & \\
 & & \delta & & 1+2\delta & & 2+\delta & & 1 \\
 & \delta & & 1+3\delta & & 3+3\delta & & 3+\delta & & 1 \\
\delta & & 1+4\delta & & 4+6\delta & & 6+4\delta & & 4+\delta & & 1 \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot,
\end{array}$$

in dem jedes Element einer Zeile durch Addition der beiden

in der vorhergehenden Zeile über ihm stehenden Elemente erhalten wird. Für $\delta = 1$ wird dieses Dreieck zum *Pascalschen.*

Wenn man als erste Diagonale die der Elemente δ ansieht, so sind die in der dritten Diagonale stehenden Elemente für $\delta = 1$ die *Dreiecks-* oder *Trigonalzahlen,* für $\delta = 2$ die *Viereckszahlen* oder *Quadratzahlen* usw., für $\delta = k - 2$ die *k-Eckszahlen.* Diese Zahlen heißen auch *Polygonalzahlen.* Der Name stammt daher, daß man ihre Einheiten in der Form eines regulären Polygons anordnen kann. Die m^{te} der k-Eckszahlen wird durch die Formel

$$\tfrac{1}{2} m\,[(k-2)\,m - (k-4)] = \binom{m}{1} + (k-2)\cdot\binom{m}{2}$$

(k Anzahl der Polygonseiten, m Seitenlänge, ausgedrückt in der Anzahl von Punkten, die auf der Seite liegen)

gegeben; die Summe der ersten m der k-Eckszahlen beträgt:

$$\tfrac{1}{6} m\,(m+1)\cdot[(m-1)\cdot(k-2) + 3].$$

Die Polygonalzahlen sollen altpythagoräischen Ursprungs sein; mit ihnen beschäftigt haben sich jedenfalls schon Hypsikles (200—100 v. Chr.) und der Neupythagoräer Nikomachus (etwa 100 n. Chr.); von Diophant, dem Vater der Arithmetik, existiert eine kleine Abhandlung über die Polygonalzahlen (deutsche Ausgabe von G. Wertheim, *die Arithmetik u. die Schrift über Polygonalzahlen des Diophantus von Alexandria,* Leipzig 1890).

Jede ganze positive Zahl läßt sich als Summe von k k-Eckszahlen darstellen. Diesen von P. Fermat (*Observations sur Diophante,* Œuvres de Fermat, publ. par Tannery (1891) **1**, 305, vgl. die oben zitierte Schrift von Wertheim, S. 162) in seinem Exemplar der Diophantausgabe (1621) von Bachet de Méziriac mit der Bemerkung „propositionem pulcherrimam et maxime generalem" ohne Beweis angegebenen Satz hat für $k = 3$ (die Dreieckszahlen) zum ersten Male Gauß in den *Disquisitiones arithmeticae,* Art. 293 (Leipzig 1801), *Ges. Werke* **1**, 348, streng bewiesen. Daß jede ganze Zahl als Summe von 4 Quadratzahlen darstellbar ist — ein wohl zuerst von Bachet de Méziriac erwähntes Theorem —, wurde von Lagrange (*Nouv. Mém. de l'acad. de Berlin* 1770, S. 123, *Œuvres* **3**, 189), dann von Euler (*Acta Acad. Petrop.,* 1777, pars II, S. 48), sowie von Gauß, a. a. O., bewiesen. Den *allgemeinen Fermatschen Satz* hat zuerst Cauchy (*Mém. de l'institut de Paris*

14, année 1813—15, S. 177, *Œuvres* (2) **6**, 320) bewiesen und ihn dabei auf folgende Weise präzisiert:

Jede ganze positive Zahl kann als Summe von k k-Eckszahlen dargestellt werden, von denen nur vier von 0 oder 1 verschiedene Werte zu haben brauchen. (Modifizierter Beweis bei Legendre, *Théorie des nombres* (1830) **2**, 340—356. Zu vgl. Bachmann, *die Arithmetik der quadratischen Formen*, Leipzig 1898, S. 154 ff.)

Die k-Eckszahlen, durch die Null ergänzt:

$$0,\ 1,\ 2+(k-2),\ 3+3(k-2),\ 4+6(k-2),\ \ldots$$

bilden eine *arithmetische Reihe zweiter Ordnung*, d. h. die Reihe ihrer Differenzen:

$$1,\ 1+(k-2),\ 1+2(k-2),\ 1+3(k-2),\ \ldots$$

ist eine arithmetische Reihe erster Ordnung, die mit 1 beginnt und die Zahl $k-2$ zur Differenz zweier Nachbarglieder hat.

Die in der *vierten* Diagonale stehenden Zahlen des oben stehenden Dreiecks sind für $\delta = 1$ die *dreiseitigen Pyramidal- oder Tetraedralzahlen*, für $\delta = 2$ die *vierseitigen Pyramidalzahlen* usw., für $\delta = k-2$ die *k-seitigen Pyramidalzahlen*. Wie die Polygonalzahlen als Summen von Punkten bei regulären Vielecken, so sind die Pyramidalzahlen geometrisch durch die Lagerung von Kugeln in Pyramiden zu finden.

Die k-seitigen Pyramidalzahlen bilden eine arithmetische Reihe dritter Ordnung, die Reihe ihrer Differenzen liefert die k-Eckszahlen.

Die m^{te} der k-seitigen Pyramidalzahlen lautet:

$$\tfrac{1}{6}m(m+1)[3+(m-1)(k-2)] = \binom{m+1}{2} + (k-2)\cdot\binom{m+1}{3}.$$

Die Summe der ersten m der k-seitigen Pyramidalzahlen beträgt:

$$\tfrac{1}{24}(m+2)(m+1)m\cdot[(m-1)(k-2)+4] = \binom{m+2}{3} + (k-2)\cdot\binom{m+2}{4}.$$

An die eingeführten Pyramidalzahlen, die von der ersten Ordnung heißen, schließen sich solche höherer Ordnung, wenn man weitergehend arithmetische Reihen vierter, fünfter usw. Ordnung einführt oder, anders ausgedrückt, die in der fünften, sechsten usw. Diagonale stehenden Zahlen des obigen Dreiecks betrachtet.

Die Binomialkoeffizienten $\binom{m+k-1}{k}$ heißen *figurierte*

Zahlen. Die für $k = 2$ sich ergebende Zahlenreihe $\binom{m+1}{2}$ liefert die Dreieckszahlen: 1, 3, 6, 10, 15, ...

Die Zahlenreihe $\binom{m+2}{3}$ stellt die dreiseitigen Pyramidalzahlen dar.

Über figurierte Zahlen vgl. Jacob Bernoullis *Ars conjectandi,* Teil II, Kap. III.

Eine Verallgemeinerung der Binomialkoeffizienten sind auch die *Polynomialkoeffizienten.* Auf sie führt die Betrachtung der positiven ganzzahligen n^{ten} Potenz eines Polynoms. Es ist:

$$(a_1 + a_2 + \cdots + a_\nu)^n = \sum \binom{n}{k_1\, k_2 \ldots k_{\nu-1}} a_1^{k_1} \cdot a_2^{k_2} \ldots a_\nu^{k_\nu}.$$

In der Summe haben $k_1, k_2, \ldots, k_\nu$ alle ganzzahligen positiven oder verschwindenden Werte zu durchlaufen, die der Bedingung $k_1 + k_2 + \cdots + k_\nu = n$ genügen. Die numerischen Koeffizienten $\binom{n}{k_1\, k_2 \ldots k_{\nu-1}}$ heißen *Polynomialkoeffizienten.* Es ist

$$\binom{n}{k_1\, k_2 \ldots k_{\nu-1}} = \frac{n!}{k_1!\, k_2! \ldots k_\nu!}.$$

Die Polynomialkoeffizienten lassen sich durch Binomialkoeffizienten ausdrücken mittels der Formel:

$$\binom{n}{k_1 k_2 \ldots k_{\nu-1}} = \binom{n}{k_1} \cdot \binom{n-k_1}{k_2} \cdot \binom{n-k_1-k_2}{k_3} \cdots \binom{n-k_1-k_2 \ldots -k_{\nu-2}}{k_{\nu-1}}.$$

§ 2. Historisches und Allgemeines über Determinanten.

Die erste Idee zu jenen Bildungen, die man heute Determinanten nennt, stammt von Leibniz (zu vgl. Cantor, *Vorl. über Gesch. der Math.*, 2. Aufl., 1901, **3**, 110). Eigentlicher Begründer der Determinantentheorie ist Gabriel Cramer, der bei dem Problem der Auflösung eines Systems linearer homogener Gleichungen in seiner *Introduction à l'analyse des lignes courbes* (Genf 1750) auf die Determinanten geführt wurde (zu vgl. Cantor **3**, 607). An Cramer knüpften Bézout, Vandermonde und Laplace an. Lagrange verwandte bei geometrischen Problemen (*Sur les pyramides* (1773), *Œuvres* **3**, 661) weitgehend Determinanten. Das Wort „determinans" ist von Gauß (*Disquisitiones arithmeticae* (1801), Sectio 5, Artikel 154 u. 267) geprägt, aber

dort nur speziell für die Determinante einer quadratischen Form von 2 und 3 Variablen verwandt worden. Eine vollständige Darstellung der Determinantensätze um ihrer selbst willen gab Cauchy in einer grundlegenden Abhandlung, *J. éc. polyt.*, Cah. **17**, 29 (1815), *Œuvres* (2) **1**, 91; er bezeichnete die von uns sogenannten Determinanten mit déterminant, später aber auch mit résultant oder fonction alternée. Allgemein üblich wurde die Bezeichnung „Determinante" durch Jacobis fundamentale Arbeit „*De formatione et proprietatibus determinantium*" (*Journ. f. Math.* **22**, 285 (1841), *Ges. Werke* **3**, 355, deutsche Ausgabe mit Anm. von Stäckel in *Ostwalds Klassikern der exakten Wiss.* No. 77).

Abgesehen von der 1851 erstmalig erschienenen Monographie von Spottiswoode, *Journ. f. Math.* **51**, 209 (1856) ist Brioschis *Teorica dei determinanti* das erste Werk über Determinanten (Pavia 1854). Wir zitieren weiter das Handbuch von Rich. Baltzer, *Theorie u. Anwendung d. Determinanten* (Leipzig 1857—1881, 1.—5. Aufl.), Trudi, *Teoria dei determinanti e loro applicazioni*, Napoli 1862, P. Gordan, *Vorl. über Invariantentheorie*, Bd. **1**: *Determinanten*, Leipzig 1885, S. Günther, *Determinanten*, 1875—77, 1.—2. Aufl., Kronecker, *Vorl. üb. d. Theorie d. Determinanten*, herausg. von Hensel, Leipzig 1903, E. Pascal, *I determinanti*, Mailand 1896, deutsche Ausg. von Leitzmann in *Teubners Sammlung* **3** (1900), R. F. Scott, *The theory of determinants and their applications*. Second ed. by Mathews 1904. Eine sehr genaue Aufstellung aller über Determinanten bis 1900 erschienenen Schriften findet man bei Th. Muir, *Quarterly Journ.* **18**, **21**, **36**, eine eingehende historische Behandlung bis zum Jahre 1841 in dem Werke desselben Verf., *The theory of determinants in the historical order of development*. London 1906. Schließlich sei noch auf den Artikel „*Determinanten*" von Netto in der *Enzyklopädie der math. Wiss.*, Bd. **1**, sowie dessen französische Bearbeitung von Vogt in der *Encyclopédie des sciences math.* verwiesen.

Irgend ein System von n^2 Größen, die in einem quadratischen Schema angeordnet sind und in ihrer Gesamtheit aufgefaßt werden sollen, nennt man eine *quadratische Matrix*. (Der Name bei A. Cayley, *Journ. f. Math.* **50**, 282 (1855), *Coll. math. papers*, Cambridge 1889, **2**, 185; der Idee nach bereits bei Euler, *Nov. Comm. Petrop.* **15**, 75 (1771).) Wir betrachten die quadratische Matrix:

$$\left\| \begin{matrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{matrix} \right\|.$$

Man bilde alle $n!$ Produkte vom Typus:

$$a_{1r_1} a_{2r_2} \dots a_{nr_n},$$

bei denen $r_1, r_2, \dots, r_n$ jede Anordnung der Zahlen 1, 2, ..., n darstellt, und lege jedem solchen Produkt das Vorzeichen + oder — bei, je nachdem die Anordnung der Indices $r_1, r_2, \dots, r_n$ zur ersten oder zweiten Klasse (zu vgl. S. 44) gehört. Die algebraische Summe der $n!$ auf diese Weise erhaltenen Produkte heißt die *Determinante der n^2 Größen a_{ik}* und wird durch das Symbol:

$$\left| \begin{matrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{matrix} \right|$$

(Cayley (1841), *Coll. math. papers* **1**, 1),

oder

$$\sum \pm a_{11} a_{22} \dots a_{nn} \text{ (Cauchy, Jacobi)}$$

oder

$$|a_{ik}| \qquad (i, k = 1, 2, \dots, n)$$

(Henry J. St. Smith (1862), *Coll. math. papers,* Oxford 1894, **1**, 229, Kronecker)

bezeichnet.

Alle Größen a der quadratischen Matrix mit dem gleichen ersten Index stehen in der nämlichen *Zeile* (*Horizontalreihe*), alle a mit dem gleichen zweiten Index in der nämlichen *Kolonne* (*Spalte, Vertikalreihe*). Der gemeinsame Name für Zeile oder Kolonne ist *Reihe.*

Die Zahl n heißt der *Grad* oder die *Ordnung der Determinante*; die Größen a_{ik} heißen ihre *Elemente.* Im besonderen nennt man $a_{11}, a_{22}, \dots, a_{nn}$ die n *Hauptelemente*, ihr Produkt $a_{11} a_{22} \dots a_{nn}$ das *Hauptglied* der Determinante; die aus den Hauptelementen bestehende Diagonale der quadratischen Matrix heißt die *Hauptdiagonale.* Zwei Elemente a_{ik} und a_{ki} der Determinante, die sich nur durch die Reihenfolge ihrer Indices unterscheiden, heißen *konjugierte Elemente.*

Die Determinante $|a_{ik}|$ *kann als Funktion der* n^2 *quadratisch geordneten Elemente* a_{ik} *durch folgende drei charakteristische Eigenschaften eindeutig definiert werden:*

1. sie ist eine ganze, lineare, homogene Funktion der Elemente jeder Zeile,

2. sie ändert nur ihr Vorzeichen, wenn man irgend zwei Zeilen miteinander vertauscht,

3. wenn $a_{11} = a_{22} = \cdots = a_{nn} = 1$ *und alle anderen Elemente 0 gesetzt werden, nimmt sie den Wert 1 an.* Statt der zweiten Eigenschaft genügt die Voraussetzung, daß die Funktion verschwinden soll, wenn irgend zwei Zeilen übereinstimmen. (Weierstraß, *Ges. Werke* **3**, 271: *Zur Determinantentheorie.* Kronecker, 17. Vorlesung. Frobenius, *Journ. f. Math.* **129**, 179 (1905).)

Die Determinanten lassen sich auch mit Hilfe der Graßmannschen Ausdehnungslehre definieren (vgl. H. Graßmann, *Ges. math. Werke* $\mathbf{I}_2$, herausg. von F. Engel, Leipzig 1896, S. 43 und 400, ferner Scott, *The theory of determinants*).

Wenn in einer Determinante alle Elemente einer Reihe Null sind, so ist die Determinante Null.

Der Wert einer Determinante bleibt ungeändert, wenn man sie umstürzt, d. h. Zeilen und Kolonnen miteinander vertauscht.

Vertauscht man in einer Determinante zwei parallele Reihen miteinander, so ändert die Determinante nur ihr Vorzeichen. Eine Determinante mit zwei identischen Parallelreihen ist Null.

Eine Determinante wird mit einer Zahl k multipliziert, indem man die Elemente einer Reihe mit k multipliziert.

Ändert man das Vorzeichen aller Elemente, für welche die Summe der Indices eine ungerade Zahl ist, so ändert sich der Wert der Determinante nicht. Multipliziert man jedes Element a_{ik} *der Determinante mit* p^{i-k}, *wobei p eine beliebige Zahl bedeutet, so behält die Determinante ihren Wert.*

Eine Determinante ist Null, wenn die Elemente einer Reihe die nämlichen Vielfachen der Elemente einer zu ihr parallelen Reihe sind.

Eine Determinante verschwindet, wenn die Elemente einer Reihe die nämlichen linearen Kombinationen der Elemente von Parallelreihen sind, und umgekehrt.

Der Wert einer Determinante ändert sich nicht, wenn man zu den Elementen einer Reihe ein und dasselbe Vielfache der entsprechenden Elemente einer Parallelreihe hinzufügt.

Eine Determinante, in der die Elemente einer Reihe Aggregate von m Gliedern sind, ist der Summe von m Determinanten gleich; diese stimmen mit der ursprünglichen Determinante völlig überein, nur in der Reihe mit den Aggregaten stehen die einzelnen Summanden.

Wenn man in einer quadratischen Matrix vom n^{ten} Grade m Zeilen und m Kolonnen wegläßt, so bleibt eine quadratische Matrix vom Grade $n - m$ übrig; die durch eine solche Matrix dargestellte Determinante heißt ein *Minor* oder eine *Subdeterminante* oder eine *Unterdeterminante* oder eine *Partialdeterminante* $n - m^{ten}$ *Grades*. Wenn ihre in der Hauptdiagonale stehenden Elemente nur Elemente aus der Hauptdiagonale der ganzen Determinante sind, so bezeichnet man sie als eine *Hauptunterdeterminante*.

Es gibt $\binom{n}{m}^2$ *Unterdeterminanten* $(n-m)^{ten}$ *Grades und ebensoviele* m^{ten} *Grades, hiervon sind* $\binom{n}{m}$ *Hauptunterdeterminanten.*

Bildet man bei einer Unterdeterminante m^{ten} Grades

$$\sum \pm a_{g_1 h_1} a_{g_2 h_2} \cdots a_{g_m h_m},$$

bei der die Ordnungszahlen der Zeilen und Kolonnen in ihrer arithmetischen Reihe aufeinanderfolgen:

$$(g_1 < g_2 < g_3 \cdots < g_m,\ h_1 < h_2 < h_3 \cdots < h_m),$$

die Summe der Indices

$$J = g_1 + g_2 + \cdots + g_m + h_1 + h_2 + \cdots + h_m,$$

so heißt diese Zahl der *Index der betreffenden Unterdeterminante* (Kroneckers *Vorl.*, S. 326); bei Trudi (*Teoria dei determinanti*, S. 17) führt sie den Namen Charakteristik.

Eine Unterdeterminante ist *gerader* oder *ungerader* Klasse, je nachdem ihr *Index* eine gerade oder ungerade Zahl ist. Zu jeder m reihigen Unterdeterminante

$$\sum \pm a_{g_1 h_1} a_{g_2 h_2} \cdots a_{g_m h_m}$$

gehört eine $n - m$ reihige

$$\sum \pm a_{g_{m+1} h_{m+1}} a_{g_{m+2} h_{m+2}} \cdots a_{g_n h_n};$$

sie wird durch Unterdrücken der Zeilen und Kolonnen, aus

denen die erste Determinante besteht, gebildet. Die Zahlen

$$g_1, g_2, \cdots, g_m, g_{m+1}, \cdots, g_n$$

und

$$h_1, h_2, \ldots, h_m, h_{m+1}, \ldots, h_n$$

$$(g_1 < g_2 < g_3 \cdots < g_m; \quad g_{m+1} < g_{m+2} \cdots < g_n;$$
$$h_1 < h_2 < h_3 \cdots < h_m; \quad h_{m+1} < h_{m+2} \cdots < h_n)$$

sind, von der Reihenfolge abgesehen, die Zahlen 1, 2, ..., n. Diese beiden Unterdeterminanten heißen *adjungiert* oder *komplementär* zueinander. Die Indices komplementärer Determinanten sind entweder beide gerade oder beide ungerade Zahlen.

Die algebraische Adjungierte oder *das algebraische Komplement* (viele Autoren verwenden auch die Bezeichnung *adjungiert* oder *komplementär im Sinne von algebraischer Adjungierten oder algebraischem Komplement*) *einer Unterdeterminante ist die zu der gegebenen Unterdeterminante komplementäre oder adjungierte mit dem positiven oder negativen Vorzeichen, je nachdem ihr Index eine gerade oder ungerade Zahl ist.*

Die algebraische Adjungierte oder das algebraische Komplement der Determinante

$$\sum \pm a_{g_1 h_1} a_{g_2 h_2} \cdots a_{g_m h_m}$$

ist demnach

$$(-1)^{\nu} \sum \pm a_{g_{m+1} h_{m+1}} a_{g_{m+2} h_{m+2}} \cdots a_{g_n h_n},$$

das heißt der Faktor von

$$a_{g_1 h_1} a_{g_2 h_2} \cdots a_{g_m h_m}$$

in der Entwicklung der ursprünglichen Determinante A.

Sieht man die n^2 Größen a_{ik} als unabhängige Variablen an, so kann die algebraische Adjungierte von

$$\sum \pm a_{g_1 h_1} a_{g_2 h_2} \cdots a_{g_m h_m}$$

in der Form eines partiellen Differentialquotienten, nämlich

$$\frac{\partial^m A}{\partial a_{g_1 h_1} \partial a_{g_2 h_2} \cdots \partial a_{g_m h_m}},$$

geschrieben werden.

Jede Determinante ist gleich der Summe der Produkte sämtlicher in m Zeilen oder Kolonnen enthaltener $\binom{n}{m}$ *m-reihiger Unterdeterminanten und ihrer algebraischen Adjungierten* (sogenannter

Zerlegungssatz von Laplace, *Recherches sur le calcul intégral et sur le système du monde, Hist. de l'acad. des sciences* 1772, *Œuvres* 8, 401, Paris 1894. Er findet sich in speziellerer Form bei Vandermonde in einer Laplaces Aufsatz unmittelbar voraufgehenden Abhandlung, *Sur l'élimination*, deutsche Ausg., Berlin 1888, S. 96). Betreffs des Laplaceschen Zerlegungssatzes vgl. auch § 11 dieses Kapitels.

Die Summe der Produkte der in m Reihen enthaltenen Unterdeterminanten und der algebraischen Adjungierten der entsprechenden Unterdeterminanten, die aus anderen m Parallelreihen entnommen sind, ist Null.

Aus diesen Sätzen folgt für $m = 1$: *Ist* A_{ik} *die algebraische Adjungierte des Elementes* a_{ik} *der Determinante A, so gelten die Relationen*:

$$a_{1i}A_{1k} + a_{2i}A_{2k} + \cdots + a_{ni}A_{nk} = \delta_{ik}A; \quad i, k = 1, 2, \ldots, n.$$
$$a_{i1}A_{k1} + a_{i2}A_{k2} + \cdots + a_{in}A_{kn} = \delta_{ik}A; \quad i, k = 1, 2, \ldots, n.$$

δ_{ik} ist das *Kroneckersche Symbol, das für* $i = k$ *den Wert 1, für* $i \gtrless k$ *den Wert 0 hat* (Kronecker, *Journ. f. Math.* **68**, 276 (1868), *Ges. Werke* **1**, 150).

Aus dem Laplaceschen Satz ergibt sich:

Eine Determinante n^{ten} *Grades, bei der alle Elemente, die m Zeilen (Kolonnen) mit* $n - m$ *Kolonnen (Zeilen) gemeinsam haben, verschwinden, ist das Produkt einer Determinante* $n - m^{ten}$ *Grades und einer* m^{ten} *Grades.*

Eine Determinante n^{ten} *Grades, bei der alle Elemente, die m Zeilen (Kolonnen) mit mehr als* $n - m$ *Kolonnen (Zeilen) gemeinsam haben, verschwinden, ist Null.*

Das Produkt zweier Determinanten gleichen Grades, von denen die eine a_{ik}, *die andere* b_{ik} *zu Elementen hat, läßt sich in Form einer Determinante gleichen Grades mit den Elementen* c_{ik} *schreiben, wobei* c_{ik} *einem der vier Ausdrücke:*

$$c_{ik} = a_{i1}b_{1k} + a_{i2}b_{2k} + \cdots + a_{in}b_{nk}$$

(Komposition der Zeilen von A mit den Kolonnen von B)

$$c_{ik} = a_{i1}b_{k1} + a_{i2}b_{k2} + \cdots + a_{in}b_{kn}$$

(Komposition der Zeilen von A mit den Zeilen von B)

$$c_{ik} = a_{1i}b_{k1} + a_{2i}b_{k2} + \cdots + a_{ni}b_{kn}$$

(Komposition der Kolonnen von A mit den Zeilen von B)

$$c_{ik} = a_{1i}b_{1k} + a_{2i}b_{2k} + \cdots + a_{ni}b_{nk}$$

Komposition der Kolonnen von A mit den Kolonnen von B)

gleich sein kann. (*Multiplikationssatz von Cauchy und Binet*, *J. éc. polyt.*, Cah. **17**, 81 u. 111 (1815), Cah. **16**, 287 (1813). Vgl. den auf die charakteristischen Eigenschaften der Determinante gegründeten Beweis in Weierstraß' Werken **3**, 283, ferner den Beweis in H. Graßmanns Werken $\mathbf{I}_2$, 400.)

Durch Hinzufügen von Zeilen und Kolonnen läßt sich der *Grad einer Determinante erhöhen.* In der Hauptdiagonale sind lauter Einser und an den noch freien Plätzen auf der einen Seite der Diagonale lauter Nullen, auf der anderen Seite beliebige Größen beizufügen.

Hat man zwei Determinanten ungleichen Grades miteinander zu multiplizieren, so bringe man die von niedrigerem Grade auf denselben Grad wie die andere.

Ein nach Zeilen und Kolonnen rechteckig geordnetes System von Elementen:

$$\begin{Vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{Vmatrix},$$

abgekürzt geschrieben:

$$\| a_{ik} \| \qquad (i = 1, 2, \ldots, m;\ k = 1, 2, \ldots, n),$$

heißt eine *rechteckige Matrix* von m Zeilen und n Kolonnen. Der erste Index der Elemente a_{ik} charakterisiert die Zeile, der zweite die Kolonne.

Die Definition einer Determinante durch charakteristische Eigenschaften läßt sich auf rechteckige Matrices erweitern. Hat man eine rechteckige Matrix von m Zeilen und n Kolonnen mit den Elementen

$$a_{ik} \qquad (i = 1, 2, \ldots, m;\ k = 1, 2, \ldots, n),$$

so ist jede Funktion der nm Elemente a_{ik}, die

1. in bezug auf die Elemente jeder Zeile der nm rechteckig geordneten Größen a_{ik} eine ganze, lineare, homogene Funktion ist, und

2. nur ihr Vorzeichen wechselt, wenn man irgend zwei Zeilen miteinander vertauscht,

eine lineare homogene Verbindung mit konstanten Koeffizienten der Determinanten m^{ten} Grades, die man aus der gegebenen Matrix bilden kann. Für $m > n$ liefert die Matrix keine Determinante m^{ten} Grades; die gesuchte Funktion ist Null. Für $m = n$ er-

hält man die mit einem von den Größen a unabhängigen Faktor multiplizierte Determinante der n^2 Größen; die Festlegung der Konstanten geschieht in diesem Fall durch die auf S. 55 erwähnte dritte Bedingung (vgl. Frobenius, *Journ. f. Math.* **129**, 179 (1905), Hensel, ebenda **126**, 73 (1903), Kronecker, 18. *Vorles.*).

Die angegebene Definition kann zum Beweis des oben erwähnten Laplaceschen Zerlegungssatzes und des erweiterten Multiplikationstheorems (siehe unten) dienen.

Hat man zwei rechteckige Matrices mit m Zeilen und n Kolonnen und stellt die Summe der Produkte der Elemente einer Zeile der einen Matrix und der entsprechenden Elemente einer Zeile der anderen Matrix her, so ergeben sich m^2 Größen. Diese bilden, wenn man sie in eine quadratische Matrix ordnet, eine Determinante m^{ten} Grades; sie heißt *das Zeilenprodukt der beiden rechteckigen Matrices.*

Erweiterter Multiplikationssatz:

Das Zeilenprodukt zweier rechteckiger Matrices von m Zeilen und n Kolonnen ist für $m < n$ gleich der Summe der Produkte der Determinanten m^{ten} Grades, die in der einen Matrix enthalten sind, in die entsprechenden Determinanten m^{ten} Grades der anderen Matrix (Cauchy u. Binet), *hingegen für $m > n$ gleich Null* (Jacobi, *de formatione et proprietatibus determinantium*, Art. 13 u. 14).

Jede Unterdeterminante einer Determinante, die das Produkt zweier Determinanten ist, kann als Zeilenprodukt zweier rechteckiger Matrices angesehen werden. Ist also

$$c_{ik} = \sum_{s=1}^{s=n} a_{is} b_{sk} \qquad (i, k = 1, 2, \ldots, n),$$

so ist die Determinante:

$$\sum \pm c_{g_1 h_1} c_{g_2 h_2} \cdots c_{g_m h_m}$$

das Zeilenprodukt der zwei rechteckigen Matrices:

$$\begin{Vmatrix} a_{g_1 1} & a_{g_1 2} & \ldots & a_{g_1 n} \\ a_{g_2 1} & a_{g_2 2} & \ldots & a_{g_2 n} \\ \vdots & & & \\ a_{g_m 1} & a_{g_m 2} & \ldots & a_{g_m n} \end{Vmatrix} \quad \text{und} \quad \begin{Vmatrix} b_{1 h_1} & b_{2 h_1} & \ldots & b_{n h_1} \\ b_{1 h_2} & b_{2 h_2} & \ldots & b_{n h_2} \\ \vdots & & & \\ b_{1 h_m} & b_{2 h_m} & \ldots & b_{n h_m} \end{Vmatrix}.$$

Die Determinante D, deren Elemente A_{ik} die algebraischen Adjungierten der Elemente a_{ik} einer gegebenen Determinante A

sind, heißt nach Cauchy (*J. éc. polyt.*, Cah. **17**, 64 (1815), *Œuvres* (2) **1**, 125) im Anschluß an Gauß' *Disquisitiones arithmeticae*, Art. 267 u. 268, die zur gegebenen *adjungierte Determinante*.

Die adjungierte Determinante einer Determinante n^{ten} Grades ist die $n-1^{te}$ Potenz der ursprünglichen (Cauchy, a. a. O., S. 82, im speziellen Fall $n=3$ bereits bei Lagrange (1773), *Œuvres* **3**, 665, sowie bei Gauß, *Disquisitiones arithmeticae*, Art. 267).

Nennt man zwei Unterdeterminanten der Determinante A und ihrer adjungierten D, die aus den gleichen Zeilen und Kolonnen von A und D gebildet sind, einander entsprechend, so läßt sich sagen:

Eine beliebige in D enthaltene Unterdeterminante D_1 vom m^{ten} Grade ist gleich der $m-1^{ten}$ Potenz der ursprünglichen Determinante, multipliziert mit derjenigen Unterdeterminante von A, die der algebraischen Adjungierten von D_1 entspricht; also:

$$\sum \pm A_{g_1 h_1} A_{g_2 h_2} \dots A_{g_m h_m} = A^{m-1} \cdot \frac{\partial^m A}{\partial a_{g_1 h_1} \partial a_{g_2 h_2} \dots \partial a_{g_m h_m}}.$$

(Vgl. § 11 dieses Kapitels. Jacobi, Formel 12 des Art. 11 *de formatione et proprietatibus determinantium*, vorher (1834) im *Journ. f. Math.* **12**, 9, *Ges. Werke* **3**, 201.)

Für $m = n-1$ folgt:

Die algebraische Adjungierte eines Elementes A_{ik} von D ist gleich a_{ik}, multipliziert mit der $n-2^{ten}$ Potenz von A.

$$\frac{\partial D}{\partial A_{ik}} = A^{n-2} \cdot a_{ik}.$$

Für $m=2$ ergibt die Jacobische Formel die häufig verwandte Relation:

$$\begin{vmatrix} A_{g_1 h_1} A_{g_1 h_2} \\ A_{g_2 h_1} A_{g_2 h_2} \end{vmatrix} = A \cdot \frac{\partial^2 A}{\partial a_{g_1 h_1} \partial a_{g_2 h_2}}.$$

Aus den Cauchy-Jacobischen Resultaten folgt:

Wenn eine Determinante den Wert Null hat, so werden auch ihre adjungierte Determinante sowie deren sämtliche Unterdeterminanten zweiten und höheren Grades Null.

Multipliziert man zwei Determinanten und ihre adjungierten Determinanten auf analoge Weise, so ist das zweite Produkt die adjungierte Determinante des ersten Produktes.

Sylvesterscher Determinantensatz (Sylvester, *Philos. Magazine* (1851), *Coll. math. papers* **1**, 243, Frobenius, *Journ. f. Math.* **86**, 53 (1879) u. **114**, 189 (1895)): Sei

$$B_{\alpha\beta} = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1m} & a_{1\beta} \\ a_{21} & a_{22} & \dots & a_{2m} & a_{2\beta} \\ \vdots & & & & \\ a_{m1} & a_{m2} & \dots & a_{mm} & a_{m\beta} \\ a_{\alpha 1} & a_{\alpha 2} & \dots & a_{\alpha m} & a_{\alpha\beta} \end{vmatrix},$$

so ist

$$\sum \pm B_{m+1,\,m+1} B_{m+2,\,m+2} \dots B_{nn} = A_m^{\,n-m-1} \cdot A,$$

wobei

$$A_m = \sum \pm a_{11} a_{22} \dots a_{mm}$$

und

$$A = \sum \pm a_{11} a_{22} \dots a_{nn}.$$

Die Determinante $B_{\alpha\beta}$ kann man als *geränderte* Determinante ansehen. Es ist

$$B_{\alpha\beta} = a_{\alpha\beta} A_m - \sum_{i=1}^{i=m} \sum_{k=1}^{k=m} A_{ik} a_{i\beta} a_{\alpha k},$$

wobei

$$A_{ik} = \frac{\partial A_m}{\partial a_{ik}} \qquad (i = 1, 2, \dots, m;\ k = 1, 2, \dots, m).$$

Hieraus folgt die häufig verwandte Formel:

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1m} & u_1 \\ a_{21} & a_{12} & \dots & a_{2m} & u_2 \\ \vdots & & & & \\ a_{m1} & a_{m2} & \dots & a_{mm} & u_m \\ v_1 & v_2 & \dots & v_m & 0 \end{vmatrix} = -\sum_{i=1}^{i=m} \sum_{k=1}^{k=m} A_{ik} u_i v_k,$$

(vgl. von Lehrbüchern etwa H. Weber, *Algebra* **1**, 95 u. 115).

Verschwinden bei einer Determinante alle Unterdeterminanten $l+1^{\text{ten}}$ Grades, hingegen nicht sämtliche l^{ten} Grades, so heißt nach Frobenius (*Journ. f. Math.* **86**, 148 (1879)) l der *Rang der Determinante.*

Man kann auch vom *Range einer Matrix*

$$\|a_{ik}\| \qquad (i = 1, 2, \dots, m;\ k = 1, 2, \dots, n)$$

sprechen. *Eine Matrix* $\|a_{ik}\|$ *ist vom Range* l, *wenn alle Determinanten* $l+1^{ten}$ *Grades, die sie enthält, und daher auch alle Determinanten höheren Grades Null sind, hingegen nicht alle Determinanten* l^{ten} *Grades verschwinden.* Damit eine Matrix $\|a_{ik}\|$ den Rang l hat, ist notwendig und hinreichend, daß sie wenigstens eine Determinante l^{ten} Grades, etwa

$$\sum \pm a_{11}\, a_{22} \cdots a_{ll},$$

enthält, die nicht verschwindet, und daß alle diejenigen Determinanten verschwinden, die sich aus

$$\sum \pm a_{11}\, a_{22} \cdots a_{ll}$$

durch Hinzufügen einer neuen Zeile und einer neuen Kolonne bilden lassen, also alle Determinanten

$$\sum \pm a_{11}\, a_{22} \cdots a_{ll}\, a_{\varrho\sigma},$$

wobei ϱ und σ nur alle Zahlen, die größer als l sind, durchlaufen (*Satz von Kronecker, Journ. f. Math.* **72**, 152 (1870), *Ges. Werke* **1**, 238).

§ 3. Symmetrische und schiefe Determinanten. Jacobische Symbole. Hermitesche Determinanten.

Wenn $a_{ik} = a_{ki}$, so heißt die Determinante *symmetrisch*; ist $a_{ik} = -a_{ki}$ $(i \gtrless k)$, so wird sie *schief* genannt (Cayley, *Journ. f. Math.* **32**, 119 (1846), *Coll. math. papers* **1**, 332) und ist schließlich $a_{ik} = -a_{ki}$ $(i \gtreqless k)$, also $a_{ii} = 0$, so heißt sie *schiefsymmetrisch* (Cayley, *Journ. f. Math.* **38**, 93 (1849), *Coll. math. papers* **1**, 410), *halbsymmetrisch* oder *alternierend*.

Das Quadrat einer jeden Determinante kann als symmetrische Determinante dargestellt werden.

In einer symmetrischen Determinante sind die algebraischen Adjungierten zweier konjugierter Elemente a_{ik} *und* a_{ki} *einander gleich; mithin ist die adjungierte Determinante einer symmetrischen Determinante ebenfalls symmetrisch.*

Wenn in einer symmetrischen Determinante alle Hauptunterdeterminanten r^{ten} *und* $r+1^{ten}$ *Grades verschwinden, so verschwinden alle Unterdeterminanten* r^{ten} *und höheren Grades* (G. Frobenius, *Journ. f. Math.* **82**, 242 (1877), S. Gundelfinger, ebenda **91**, 229 (1881), G. Frobenius **114**, 192 (1895)).

Eine schiefsymmetrische Determinante von ungeradem Grade ist Null; ihre adjungierte Determinante ist symmetrisch.

Eine schiefsymmetrische Determinante geraden Grades ist das vollständige Quadrat einer ganzen rationalen Funktion ihrer Elemente a_{ik} (Cayley, *Journ. f. Math.* **38**, 93 (1849), *Coll. math. papers* **1**, 410), vgl. auch Mertens, *Journ. f. Math.* **82**, 207 (1877); *die adjungierte Determinante einer schiefsymmetrischen Determinante geraden Grades ist ebenfalls schiefsymmetrisch.*

Der Ausdruck, dessen Quadrat die schiefsymmetrische Determinante geraden Grades ist und der bereits von Jacobi, *Journ. f. Math.* **2**, 356 (1827) u. ebenda **29**, 237 (1845), *Ges. Werke* **4**, 26 u. 420 behandelt wurde, heißt nach Cayley die *Jacobische* (*Journ. f. Math.* **38**, 94 (1849)) oder *Pfaffsche* (*Journ. f. Math.* **50**, 300 (1855)) *Funktion* oder nach Scheibner (*Leipz. Ber.* (1859), 151) *Halbdeterminante.*

Die Anzahl der Glieder einer Jacobischen Funktion, deren Quadrat eine schiefsymmetrische Determinante n^{ten} Grades ist, beträgt $1 \cdot 3 \cdot 5 \cdots (n-3)(n-1)$. Jedes Glied ist ein Produkt aus $\frac{n}{2}$ Faktoren, deren Indices zusammen sämtliche Zahlen der Reihe $1, 2, \ldots, n$ sind. Unter den Gliedern befindet sich auch das Glied $a_{12}\, a_{34}\, a_{56} \cdots a_{n-1,n}$; wählt man dieses Glied mit dem positiven Vorzeichen, so ist die Quadratwurzel aus der schiefsymmetrischen Determinante eindeutig bestimmt. Den auf diese Weise festgelegten Jacobischen oder Pfaffschen Ausdruck bezeichnet man nach Jacobi durch *das Symbol* $(1\,2\,3 \ldots n)$. *Das Symbol* $(1\,2\,3 \ldots n)$ *berechnet sich rekurrent mittels der Formel:*

$$(1\,2 \ldots n) = (1\,2)(3\,4 \ldots n) + (1\,3)(4\,5 \ldots n\,2) + \cdots$$
$$+ (1\,k)(k+1\,k+2 \ldots n\,2 \ldots k-1) + \cdots (1\,n)(2\,3 \ldots n-1);$$

hierbei bedeuten $(1\,2), (1\,3), \ldots, (1\,n)$ die Elemente $a_{12}, a_{13}, \ldots, a_{1n}$.

Das Vorzeichen eines Jacobischen Symbols ändert sich, wenn man zwei Elemente miteinander vertauscht.

Ist in einer schiefsymmetrischen Determinante m der höchste Grad nichtverschwindender Unterdeterminanten (d. h. die Determinante ist vom Range m), *so ist m notwendig eine gerade Zahl und unter den nichtverschwindenden Unterdeterminanten* m^{ten} *Grades befinden sich auch Hauptunterdeterminanten.*

Wenn in einer schiefsymmetrischen Determinante alle Hauptunterdeterminanten $2r^{\text{ten}}$ *Grades Null sind, so verschwinden auch alle Unterdeterminanten* $2r - 1^{\text{ten}}$ *Grades* (Frobenius, *Journ. f. Math.* **82**, 242 (1877)). Vgl. hierzu auch die Anmerkungen von

F. Engel in H. Graßmanns *Ges. math. Werken* $\mathbf{1}_2$, 490 sowie E. v. Weber, *Vorl. üb. das Pfaffsche Problem, Teubners Samml.* **2**, 19 ff.

Jede schiefe Determinante, deren Hauptelemente gleich 1 sind, ist eine Summe von Quadraten (Cayley, *Journ. f. Math.* **38**, 96 (1849), *Coll. math. papers* **1**, 413).

Eine Determinante, bei der die Elemente a_{ik} und a_{ki} konjugiert imaginär, die Elemente a_{ii} im besonderen reell sind, heißt eine *Hermitesche Determinante* (Hermite, *Journ. f. Math.* **47**, 345 (1854), **52**, 39 (1856) u. **53**, 183 (1857), *C. R.* **41** (1855), 181). Sie hat stets einen reellen Wert.

Bei einer Hermiteschen Determinante sind die algebraischen Adjungierten zweier konjugierter Elemente a_{ik} und a_{ki} konjugiert imaginär; mithin ist die adjungierte Determinante einer Hermiteschen Determinante wieder eine Hermitesche Determinante.

Die Gleichung:

$$\begin{vmatrix} a_{11}-x & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22}-x & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn}-x \end{vmatrix} = 0$$

hat, wenn die Determinante der a_{ik} eine Hermitesche ist, nur reelle Wurzeln (Hermite, *C. R.* **41**, 181—183); für jede m fache Wurzel der Gleichung verschwinden auch alle Unterdeterminanten $n-1^{\text{ter}}$, $n-2^{\text{ter}}$, ..., $n-m+1^{\text{ter}}$ Ordnung der linksstehenden Determinante (Clebsch, *Journ. f. Math.* **57**, 327 (1860), ebenda **62**, 232 (1863), Christoffel, ebenda **63**, 255 (1864)).

Nimmt man die Elemente einer Hermiteschen Determinante reell an, so entsteht eine *reelle symmetrische Determinante.* Dieser besondere Fall ergibt den für allgemeines n zuerst von Cauchy (*Exerc. de math.* **4**, 140, *Œuvres* (2) **9**, 174), für $n=3$ schon 1773 von Lagrange (*Œuvres* **3**, 605) aufgestellten Satz:

Die Gleichung:

$$\begin{vmatrix} a_{11}-x & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22}-x & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn}-x \end{vmatrix} = 0$$

hat, wenn die Determinante der a symmetrisch ist und die a reelle Größen sind, nur reelle Wurzeln; man nennt diese Gleichung die Säkulargleichung. Von den mehrfachen Wurzeln hat

Weierstraß (*Monatsb. d. Berl. Akad.* (1858) 213, (1868) 336, (1879) 430, *Ges. Werke* **1**, 238, **2**, 42, **3**, 139) zuerst bewiesen: *Für jede m fache Wurzel der Gleichung verschwinden alle Unterdeterminanten* $n - 1^{ter}$, $n - 2^{ter}$, *bis* $n - m + 1^{ter}$ *Ordnung der linksstehenden Determinante.*

Die Säkulargleichung ist für $n = 3$ bei der Bestimmung der Hauptachsen der Flächen 2ter Ordnung wichtig. Literatur über diese Gleichung findet man bei Clebsch-Lindemann, *Vorl. über Geometrie* **2**, 168, Leipzig 1897. Vgl. ferner die im § 9 dieses Kapitels über Scharen reeller quadratischer und Hermitescher Formen angegebenen Resultate.

Nimmt man die Elemente einer Hermiteschen Determinante rein imaginär an und unterdrückt den Faktor i, so entsteht eine *reelle schiefsymmetrische Determinante.* Der für eine Hermitesche Determinante ausgesprochene Satz drückt sich in diesem besonderen Fall folgendermaßen aus.

Die Gleichung:

$$\begin{vmatrix} -x & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & -x & a_{23} & \dots & a_{2n} \\ \vdots & & & & \\ a_{n1} & a_{n2} & a_{n3} & \dots & -x \end{vmatrix} = 0$$

hat, wenn die Determinante der a schiefsymmetrisch ist und die a reelle Größen sind, nur rein imaginäre Wurzeln; für jede m fache Wurzel der Gleichung verschwinden auch alle Unterdeterminanten $n - 1^{ter}$, $n - 2^{ter}$, $\dots$, $n - m + 1^{ter}$ *Ordnung der linksstehenden Determinante.*

§ 4. Spezielle Determinanten.

Eine von Hankel näher untersuchte Determinante (Hankel, *Diss.*, Göttingen 1861), auf die Jacobi schon 1835 im *Journ. f. Math.* **15**, kam (zu vgl. Kronecker, *Ges. Werke* **2**, 105), wird auf folgende Art gebildet:

$$P = \begin{vmatrix} a_0 & a_1 & a_2 & \dots & a_{n-1} \\ a_1 & a_2 & a_3 & \dots & a_n \\ \vdots & & & & \\ a_{n-1} & a_n & a_{n+1} & \dots & a_{2n-2} \end{vmatrix};$$

das Element a_{ik} ist gleich a_{i+k-2}, also nur von der Summe der Indices abhängig. Man nennt diese Determinante *orthosymmetrisch* (Hankel), *persymmetrisch* (Sylvester, *Philos.*

Trans. (1853), *Coll. math. papers*, Cambridge 1904, **1**, 448 (p. 584) und in neuerer Bezeichnung *rekurrierend* (Frobenius, *Journ. f. Math.* **114**, 200 (1894)).

Bildet man die Differenzen:

$$\Delta_0^{(0)} = a_0,$$

$$\Delta_1^{(1)} = a_1 - a_0,$$

$$\Delta_2^{(1)} = a_2 - a_1,\ \Delta_2^{(2)} = \Delta_2^{(1)} - \Delta_1^{(1)},$$

$$\Delta_k^{(1)} = a_k - a_{k-1},\ \Delta_k^{(2)} = \Delta_k^{(1)} - \Delta_{k-1}^{(1)},\ \Delta_k^{(3)} = \Delta_k^{(2)} - \Delta_{k-1}^{(2)},\ \ldots,$$

so erhält man

$$P = \begin{vmatrix} \Delta_0^{(0)} & \Delta_1^{(1)} & \Delta_2^{(2)} & \Delta_3^{(3)} & \ldots & \Delta_{n-1}^{(n-1)} \\ \Delta_1^{(1)} & \Delta_2^{(2)} & \Delta_3^{(3)} & & \ldots & \Delta_n^{(n)} \\ \Delta_2^{(2)} & \Delta_3^{(3)} & & & \ldots & \Delta_{n+1}^{(n+1)} \\ \vdots & & & & & \\ \Delta_{n-1}^{(n-1)} & \Delta_n^{(n)} & \Delta_{n+1}^{(n+1)} & & \ldots & \Delta_{2n-2}^{(2n-2)} \end{vmatrix};$$

die rekurrierende Determinante der Größen $a_0, a_1, \ldots, a_{2n-2}$ *ist also gleich der rekurrierenden Determinante der Differenzen*

$$\Delta_0^{(0)},\ \Delta_1^{(1)},\ \Delta_2^{(2)},\ \ldots,\ \Delta_{2n-2}^{(2n-2)}.$$

Sind die Differenzen n^{ter} *Ordnung Null, so wird*

$$P = (-1)^{\frac{n(n-1)}{2}} [\Delta_{n-1}^{(n-1)}]^n;$$

sind schon die Differenzen niedrigerer als n^{ter} *Ordnung Null, so wird* $P = 0$.

Rekurrierende Determinanten spielen in der Invariantentheorie (vgl. Pascal, *die Determinanten*, S. 70), beim Sturmschen Theorem und der Diskriminantenbildung (vgl. das Kapitel über *Algebra*) eine wichtige Rolle.

Eine spezielle Form der rekurrierenden Determinante ist die *zyklische Determinante*. Sie hat die Gestalt:

$$A = \begin{vmatrix} a_1 & a_2 & \ldots & a_n \\ a_2 & a_3 & \ldots & a_1 \\ a_3 & a_4 & \ldots & a_2 \\ \vdots & & & \\ a_n & a_1 & \ldots & a_{n-1} \end{vmatrix}$$

Jede zyklische Determinante vom n^{ten} Grade läßt sich u. ein Produkt von n Faktoren, die von den *n Elementen a rational abhängen, zerlegen, nämlich*

$$A = (-1)^{\frac{(n-1)(n-2)}{2}} \varphi(\alpha_1)\varphi(\alpha_2)\dots\varphi(\alpha_n);$$

hierbei ist:

$$\varphi(z) = a_1 + a_2 z + a_3 z^2 + \dots + a_n z^{n-1}$$

und $\alpha_1, \alpha_2, \dots, \alpha_n$ die Wurzeln von $z^n = 1$.

Anstatt der zyklischen Determinante betrachtet man häufig die aus ihr durch Zeilenvertauschung hervorgehende, bei der in der ersten Kolonne die Elemente

$$a_1, a_n, a_{n-1}, a_{n-2}, \dots, a_2$$

stehen; sie ist in bezug auf die Nebendiagonale symmetrisch. Man nennt diese von der obigen nur durch das Vorzeichen

$$(-1)^{\frac{(n-1)(n-2)}{2}}$$

verschiedene Determinante eine *Zirkulante*. Literatur in Pascals Determinanten, § 20.

Die Vandermondesche oder Cauchysche Determinante wird durch:

$$\Delta(a_1, a_2, \dots, a_n) = \begin{vmatrix} 1 & 1 & 1 & \dots & 1 \\ a_1 & a_2 & a_3 & \dots & a_n \\ a_1^2 & a_2^2 & a_3^2 & \dots & a_n^2 \\ \vdots & & & & \\ a_1^{n-1} & a_2^{n-1} & a_3^{n-1} & \dots & a_n^{n-1} \end{vmatrix}$$

dargestellt. Für $n = 3$ ist sie von Vandermonde (1770), *Résolution des équations*, deutsche Ausg., Berlin 1888, S. 8 eingeführt und dann für allgemeines n von Cauchy, *J. éc. polyt.*, Cah. **17**, S. 48, *Œuvres* (2) **1**, 109, untersucht worden. Vgl. auch Jacobi, *de functionibus alternantibus*, *Journ. f. Math.* **22**, 360 (1841), *Ges. Werke* **3**, 439, deutsche Ausgabe von Staeckel in *Ostwalds Klassikern der exakten Wissenschaften* No. 77, S. 50.

Δ ist gleich dem *Differenzenprodukt von* $\frac{n(n-1)}{2}$ *Faktoren:*

$$\Delta = \prod_{i,j} (a_j - a_i) \quad (j > i).$$

Δ ist eine *alternierende* Funktion der Größen $a_1, a_2, \ldots, a_n$. Eine Funktion von n Größen $a_1, a_2, \ldots, a_n$ heißt alternierend (Cauchy, a. a. O., S. 30), wenn sie bei allen Vertauschungen der n Größen höchstens ihr Vorzeichen ändert. Die Cauchysche Determinante ist für die *Diskriminante* einer algebraischen Gleichung (vgl. Algebra) von größter Wichtigkeit.

Das *Quadrat einer Cauchyschen Determinante* ist gleich

$$\begin{vmatrix} s_0 & s_1 & s_2 & \ldots & s_{n-1} \\ s_1 & s_2 & s_3 & \ldots & s_n \\ s_2 & s_3 & s_4 & \ldots & s_{n+1} \\ \vdots & & & & \\ s_{n-1} & s_n & s_{n+1} & \ldots & s_{2n-2} \end{vmatrix},$$

also eine *rekurrierende* Determinante, wobei

$$s_i = a_1^i + a_2^i + \cdots + a_n^i.$$

Die *Näherungsbrüche des Kettenbruches*

$$a_1 + \cfrac{b_2}{a_2 + \cfrac{b_3}{a_3 + \cdots}}$$

führen auf die *sogenannte Kettenbruchdeterminante*

$$A_n = \begin{vmatrix} a_1 & b_2 & 0 & 0 & 0 & \ldots & 0 \\ -1 & a_2 & b_3 & 0 & 0 & \ldots & 0 \\ 0 & -1 & a_3 & b_4 & 0 & \ldots & 0 \\ 0 & 0 & -1 & a_4 & b_5 & \ldots & 0 \\ \vdots & & & & & & \\ 0 & 0 & 0 & \ldots & & -1 & a_n \end{vmatrix}.$$

Der Kettenbruch

$$a_1 + \cfrac{b_2}{a_2 + \cdots \cfrac{b_n}{a_n}}$$

ist gleich A_n dividiert durch das algebraische Komplement des Elementes a_1 in der Determinante A_n; sein Wert ist also $\dfrac{A_n}{\dfrac{\partial A_n}{\partial a_1}}$.

Definiert man $A_0 = +1$ und $A_1 = a_1$, so bestehen die

Rekursionsformeln

$$A_\nu = A_{\nu-1}\, a_\nu + b_\nu A_{\nu-2} \qquad (\nu = 2, 3, 4, \ldots).$$

Betreffs Literatur sei verwiesen auf Pascal, *Determinanten*, S. 156 und Günther, *Determinanten*, 2. Aufl., S. 123, sowie vorzüglich auf Th. Muirs bis 1880 gehende historische Behandlung der Kontinuanten (*Edinburgh R. S. Proc.* **25**, 129, 648). Die Kettenbruchdeterminanten heißen auch *Kontinuanten* (Muir, a. a. O., S. 668); für $b_2 = b_3 = \cdots b_n = 1$ erhält man die sogenannten *einfachen Kontinuanten*, mit denen die Theorie bei Sylvester (*Philos. Magazine* (1853), *Coll. math. papers* **1**, 616) anfängt. Mit einfachen Kontinuanten und ihrer Verwendung für Zahlentheorie beschäftigt sich die *Straßburger Dissertation* von R. E. Moritz (1902).

Wir geben noch die Werte einiger *arithmetischer* Determinanten, d. h. solcher, die aus besonderen Zahlen gebildet sind und besondere Zahlenwerte haben.

Die aus Binomialkoeffizienten gebildete Determinante

$$\begin{vmatrix} 1 & 1 & \cdots & 1 \\ \binom{c+m}{1} & \binom{c+m+1}{1} & \cdots & \binom{c+2m}{1} \\ \binom{c+m+1}{2} & \binom{c+m+2}{2} & \cdots & \binom{c+2m+1}{2} \\ \vdots & & & \\ \binom{c+2m-1}{m} & \binom{c+2m}{m} & \cdots & \binom{c+3m-1}{m} \end{vmatrix}$$

hat, unabhängig von c und m, den Wert 1 (zu vgl. Baltzer, *Determinanten*, 5. Aufl., S. 24).

Die Zeipelsche Determinante

$$\begin{vmatrix} \binom{m}{p} & \binom{m}{p+1} & \cdots & \binom{m}{p+r} \\ \binom{m+1}{p} & \binom{m+1}{p+1} & \cdots & \binom{m+1}{p+r} \\ \vdots & & & \\ \binom{m+r}{p} & \binom{m+r}{p+1} & \cdots & \binom{m+r}{p+r} \end{vmatrix}$$

ist gleich

$$\frac{\binom{m+r}{r+1}\binom{m+r-1}{r+1}\cdots\binom{m+r-p+1}{r+1}}{\binom{p+r}{r+1}\binom{p+r-1}{r+1}\cdots\binom{r+1}{r+1}},$$

(Zu vgl. Günther, *Determinanten*, S. 80 und Pascal, *Determinanten*, S. 133, sowie Netto, *Lehrb. der Kombinatorik*, S. 256).

Die Sternsche Determinante (*Journ. f. Math.* **66**, 285 (1866))

$$\begin{vmatrix} 1 & 1 & \dots & 1 \\ \binom{x_1}{1} & \binom{x_2}{1} & & \binom{x_n}{1} \\ \binom{x_1}{2} & \binom{x_2}{2} & \dots & \binom{x_n}{2} \\ \vdots & & & \\ \binom{x_1}{n-1} & \binom{x_2}{n-1} & \dots & \binom{x_n}{n-1} \end{vmatrix}$$

ist gleich

$$\frac{\Delta}{2^{n-2}\,3^{n-3}\,4^{n-4}\dots(n-1)},$$

wobei Δ *die Vandermondesche oder Cauchysche Determinante der Größen* $x_1, x_2, \dots, x_n$ *bedeutet.*

Die Smithsche Determinante (Henry J. St. Smith, *Coll. math. papers* **2**, 161, Oxford)

$$\begin{vmatrix} (1,1) & (1,2) & \dots & (1,n) \\ \vdots & & & \\ (n,1) & (n,2) & \dots & (n,n) \end{vmatrix},$$

in der unter (i, j) *der größte gemeinsame Teiler der beiden ganzen positiven Zahlen* i, j *verstanden wird, ist gleich*

$$\varphi(1)\,\varphi(2)\dots\varphi(n),$$

worin $\varphi(k)$ *die Anzahl der Zahlen bedeutet, die kleiner als* k *und zu* k *relativ prim sind.*

Determinanten der Form:

$$R_n = \begin{vmatrix} a_{10} & -t & 0 & 0 & \dots & 0 \\ a_{20} & a_{21} & -t & 0 & \dots & 0 \\ a_{30} & a_{31} & a_{32} & -t & \dots & 0 \\ \vdots & & & & & \\ a_{n0} & a_{n1} & a_{n2} & a_{n3} & \dots & a_{n,n-1} \end{vmatrix}$$

sind von E. Pascal (*Rend. Ist. Lomb.* (2) **40**, 293 (1907)) studiert worden. Die *Eulerschen*, die *Bernoullischen* (vgl. Kap. XX) und andere von E. Pascal (*Rend. Ist. Lomb.* (2) **40**, 461) ein-

geführte Zahlen lassen sich als spezielle Determinanten R_n, d. aus Binomialkoeffizienten gebildet sind, darstellen.

Für derartige Determinanten gelten die Relationen:

$$R_n = a_{n0}\,t^{n-1} + a_{n1}R_1\,t^{n-2} + \cdots + a_{n,n-1}R_{n-1},$$
$$R_n = A_1^{(n)}\,t^{n-1} + A_2^{(n)}\,t^{n-2} + \cdots + A_n^{(n)},$$

wobei

$$A_r^{(s)} = \sum_i a_{s\,i_1}\,a_{i_1\,i_2}\ldots a_{i_{r-1}\,0}$$

ist und die Summe über alle möglichen Werte der Zahlen i, die den Ungleichungen:

$$s > i_1 > i_2 > i_3 \cdots > 0$$

genügen, zu erstrecken ist.

Bedeuten $t_1, t_2, \ldots, t_n$ willkürliche Größen und multipliziert man jedes Element a_{ik} einer Determinante R_n mit $t_{k+1}t_{k+2}\ldots t_i$, so wird hierdurch R_n selbst mit $t_1 t_2 \ldots t_n$ multipliziert.

§ 5. Systeme linearer Gleichungen.

Gegeben sei das System von m linearen Gleichungen mit den n Unbekannten $x_1, x_2, \ldots, x_n$:

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= y_1,\\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= y_2,\\
&\vdots\\
a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= y_m.
\end{aligned}$$

Die Matrix

$$\| a_{ik} \| \qquad (i = 1, 2, \ldots, m;\ k = 1, 2, \ldots, n)$$

soll *die Matrix der Koeffizienten* heißen und *den Rang l besitzen* (vgl. S. 62). Eine der nicht verschwindenden Determinanten l^{ten} Grades, die in der Matrix $\| a_{ik} \|$ enthalten sind, sei die *Determinante*[1]):

$$A = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1l} \\ a_{21} & a_{22} & \cdots & a_{2l} \\ \vdots & & & \\ a_{l1} & a_{l2} & \cdots & a_{ll} \end{vmatrix}.$$

1) Durch geeignete Bezeichnung der Variablen und entsprechende Anordnung der vorgelegten Gleichungen ist es stets zu erzielen, daß eine der nicht verschwindenden Determinanten l^{ten} Grades die Determinante A wird.

Seiner Natur nach ist *l jedenfalls nicht größer als die kleinere der Zahlen m und n.*

Damit die gegebenen Gleichungen endliche Lösungen besitzen, also sich nicht widersprechen, ist im Falle $l < m$ das Verschwinden aller Determinanten:

$$\Gamma_\varrho = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1l} & y_1 \\ a_{21} & a_{22} & \dots & a_{2l} & y_2 \\ \vdots & & & & \\ a_{l1} & a_{l2} & \dots & a_{ll} & y_l \\ a_{\varrho 1} & a_{\varrho 2} & \dots & a_{\varrho l} & y_\varrho \end{vmatrix}$$

notwendig und hinreichend; hierbei hat ϱ alle Zahlen, die größer als l sind, zu durchlaufen. Für $l = m$ haben die gegebenen Gleichungen stets endliche Lösungen.

Diesem Theorem kann man nach dem Kroneckerschen Satz (S. 63) folgende gleichwertige Fassung geben:

Damit die gegebenen Gleichungen miteinander vereinbar sind, d. h. endliche Lösungen zulassen, ist notwendig und hinreichend, daß die Matrix:

$$\left\| \begin{matrix} a_{11} & a_{12} & \dots & a_{1n} & y_1 \\ a_{21} & a_{22} & \dots & a_{2n} & y_2 \\ \vdots & & & \vdots & \\ a_{m1} & a_{m2} & \dots & a_{mn} & y_m \end{matrix} \right\|$$

denselben Rang l hat, wie die Matrix der Koeffizienten:

$$\left\| \begin{matrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{matrix} \right\|$$

(Frobenius, *Journ. f. Math.* **86**, 171 (1879), Satz IV).

Ist l der gemeinsame Rang der zwei angegebenen Matrices, so reduziert sich das System der m gegebenen Gleichungen auf das der ersten l von ihnen. Die Lösungen des Systems bilden eine $n - l$-fache Mannigfaltigkeit; l der Unbekannten lassen sich nämlich als lineare Funktionen der übrigen $n - l$, die völlig willkürliche Werte erhalten können, ausdrücken.

Die Werte von $x_1, x_2, \dots, x_l$ werden durch die Formel:

$$x_i = \frac{\Delta_i}{A} \qquad (i = 1, 2, \dots l)$$

gegeben; dabei bedeutet Δ_i die aus A hervorgehende Determinante, wenn man die Größen der i^{ten} Kolonne von A durch:

$$\begin{aligned} y_1' &= y_1 - a_{1,l+1}x_{l+1} - a_{1,l+2}x_{l+2} \cdots - a_{1n}x_n, \\ y_2' &= y_2 - a_{2,l+1}x_{l+1} - a_{2,l+2}x_{l+2} \cdots - a_{2n}x_n, \\ &\vdots \\ y_l' &= y_l - a_{l,l+1}x_{l+1} - a_{l,l+2}x_{l+2} \cdots - a_{ln}x_n, \end{aligned}$$

ersetzt. Es ist also:

$$x_i = \frac{\Delta_i}{A} = \frac{A_{1i}y_1' + A_{2i}y_2' + \cdots + A_{li}y_l'}{A} \qquad (i = 1, 2, \ldots, l);$$

hierbei bedeutet A_{ik} die algebraische Adjungierte des Elementes a_{ik} in der Determinante A, d. h.

$$A = a_{1i}A_{1i} + a_{2i}A_{2i} + \cdots + a_{li}A_{li}.$$

Den Größen $x_{l+1}, x_{l+2}, \ldots, x_n$ können beliebige Werte beigelegt werden.

Ist im besonderen der gemeinsame Rang l gleich der Anzahl n der Unbekannten, so reduzieren sich die gegebenen Gleichungen auf ein System von n Gleichungen mit n Unbekannten $x_1, x_2, \ldots, x_n$ und nicht verschwindender Determinante A:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n - y_1 &= 0, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n - y_2 &= 0, \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n - y_n &= 0. \end{aligned}$$

Die y' werden die y selbst, und es gibt nur ein einziges aus den Gleichungskoeffizienten und den bekannten Termen rational gebildetes Wertsystem für $x_1, x_2, \ldots, x_n$, das den Gleichungen genügt:

$$x_i = \frac{A_{1i}y_1 + A_{2i}y_2 + \cdots + A_{ni}y_n}{A} \qquad (i = 1, 2, \ldots, n),$$

(Formel von Gabriel Cramer, *Introduction à l'analyse des lignes courbes* 1750, Append., S. 657). In diesem Falle kann man sich die Lösung durch folgende Regel merken: Man schreibe die Matrix

$$\left\| \begin{matrix} a_{11} & a_{12} & \cdots & a_{1n} & -y_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & -y_2 \\ \vdots & & & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} & -y_n \end{matrix} \right\|.$$

hin; $x_1 : x_2 : \cdots : x_n : 1$ verhalten sich alsdann wie die mit abwechselnden Vorzeichen versehenen Determinanten, die man aus der angegebenen Matrix durch sukzessives Streichen der ersten, zweiten, usw. schließlich letzten Kolonne erhält.

Lineare homogene Gleichungen. Sind $y_1 = y_2 = \cdots = y_m = 0$, so hat man ein System von *m linearen homogenen Gleichungen:*

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = 0 \qquad (i = 1, 2, \ldots, m)$$

mit n Unbekannten. Soll ein solches System eine von der selbstverständlichen Lösung $x_1 = x_2 = \cdots = x_n = 0$ *verschiedene Lösung zulassen, so ist hierzu nur notwendig, daß die Matrix der Koeffizienten:*

$$\left\| \begin{matrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{matrix} \right\|$$

den Rang $l < n$ *hat.*

Sind

$$\begin{matrix} b_1^{(1)}, b_2^{(1)}, \ldots, b_n^{(1)}, \\ b_1^{(2)}, b_2^{(2)}, \ldots, b_n^{(2)}, \\ \vdots \\ b_1^{(\varrho)}, b_2^{(\varrho)}, \ldots, b_n^{(\varrho)}, \end{matrix}$$

irgend ϱ *Systeme von partikulären Lösungen der homogenen Gleichungen, d. h. wird das Gleichungssystem erfüllt, wenn man* $b_1^{(\alpha)}, b_2^{(\alpha)}, \ldots, b_n^{(\alpha)}$ *für* $x_1, x_2, \ldots, x_n$ *setzt, so stellen auch:*

$$k_1 b_1^{(1)} + k_2 b_1^{(2)} + \cdots + k_\varrho b_1^{(\varrho)}, \quad k_1 b_2^{(1)} + k_2 b_2^{(2)} + \cdots + k_\varrho b_2^{(\varrho)},$$
$$\cdots, k_1 b_n^{(1)} + k_2 b_n^{(2)} + \cdots k_\varrho b_n^{(\varrho)},$$

wo $k_1, k_2, \ldots, k_\varrho$ *willkürliche Konstanten bedeuten, ein Lösungssystem dar.*

Irgend $\varrho \leqq n$ *Systeme partikulärer Lösungen heißen unabhängig, wenn es keine Konstanten* $c_1, c_2, \ldots, c_\varrho$ *gibt, welche den n Gleichungen:*

$$c_1 b_\lambda^{(1)} + c_2 b_\lambda^{(2)} + c_\varrho b_\lambda^{(\varrho)} = 0 \qquad (\lambda = 1, 2, \ldots, n)$$

genügen, ohne daß alle ϱ *Größen* $c_1, c_2, \ldots, c_\varrho$ *gleichzeitig ver-*

schwinden. ϱ Lösungssysteme sind dann und nur dann linear unabhängig, wenn der Rang der Matrix:

$$\left\| \begin{matrix} b_1^{(1)} & b_2^{(1)} & \dots & b_n^{(1)} \\ b_1^{(2)} & b_2^{(2)} & \dots & b_n^{(2)} \\ \vdots & \vdots & & \vdots \\ b_1^{(\varrho)} & b_2^{(\varrho)} & \dots & b_n^{(\varrho)} \end{matrix} \right\|$$

genau gleich ϱ ist.

Sind die ϱ Lösungssysteme unabhängig, so sind die ϱ linearen homogenen Funktionen:

$$b_1^{(\alpha)} x_1 + b_2^{(\alpha)} x_2 + \dots + b_n^{(\alpha)} x_n \qquad (\alpha = 1, 2, \dots \varrho)$$

nicht in linearer Dependenz und umgekehrt (vgl. S. 83).

Hat die Matrix

$$\| a_{ik} \| \qquad (i = 1, 2, \dots, m;\ k = 1, 2, \dots, n)$$

des homogenen Gleichungssystemes den Rang $l < n$, so besitzt das Gleichungssystem genau $n - l$ unabhängige Systeme partikulärer Lösungen

$$b_1^{(\alpha)},\ b_2^{(\alpha)},\ \dots,\ b_n^{(\alpha)} \qquad (\alpha = 1, 2, \dots, n - l).$$

Diese kann man aus l geeigneten Gleichungen des Systemes finden; die übrigen $m - l$ Gleichungen kann man einfach fortlassen.

Die allgemeinste Lösung lautet:

$$k_1 b_\tau^{(1)} + k_2 b_\tau^{(2)} + \dots + k_{n-l} b_\tau^{(n-l)} \qquad (\tau = 1, 2, \dots, n),$$

wobei $k_1, k_2, \dots, k_{n-l}$ willkürliche Konstanten bedeuten. Aus der allgemeinen Lösung erhält man jede Lösung, indem man $k_1, k_2, \dots, k_{n-l}$ bestimmte Werte erteilt.

Ist l der Rang der Matrix

$$\| a_{ik} \| \qquad (i = 1, 2, \dots, m;\ k = 1, 2, \dots, n)$$

und verschwindet die Determinante

$$\sum \pm a_{11}\, a_{22} \dots a_{ll}$$

nicht, so *findet man $n - l$ Systeme unabhängiger Lösungen auf* folgende Weise: Man wähle willkürlich $n(n - l)$ Größen

$$U_\alpha^{(\beta)} \qquad (\alpha = 1, 2, \dots, n;\ \beta = 1, 2, \dots, n - l),$$

die nur der Bedingung zu genügen haben, daß die Determinante:

$$D = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{l1} & a_{l2} & \dots & a_{ln} \\ U_1^{(1)} & U_2^{(1)} & \dots & U_n^{(1)} \\ U_1^{(2)} & U_2^{(2)} & \dots & U_n^{(2)} \\ \vdots & \vdots & & \vdots \\ U_1^{(n-l)} & U_2^{(n-l)} & \dots & U_n^{(n-l)} \end{vmatrix}$$

einen von Null verschiedenen Wert hat. $n(n-l)$ derartige Größen $U_\alpha^{(\beta)}$ lassen sich auf unendlich viele Weisen finden, z. B. indem man

$$U_{l+\beta}^{(\beta)} = 1 \qquad (\beta = 1, 2, \dots, n-l)$$

und alle anderen Größen $U_\alpha^{(\beta)} = 0$ wählt. Sind α_{ik} die algebraischen Adjungierten der Elemente der Determinante D, ist also

$$\sum \pm \alpha_{11}\, \alpha_{22} \dots \alpha_{nn}$$

die adjungierte Determinante von D, so bilden die $n(n-l)$ Elemente der $n-l$ letzten Horizontalreihen der adjungierten Determinante von D $n-l$ Systeme unabhängiger Lösungen des homogenen Gleichungssystems. Man hat:

$$\alpha_{k1}a_{i1} + \alpha_{k2}a_{i2} + \dots + \alpha_{kn}a_{in} = 0 \qquad (k = l+1, l+2, \dots, n;\ i = 1, 2, \dots, m)$$

und die Matrix

$$\| \alpha_{kt} \| \qquad (k = l+1, l+2, \dots, n;\ i = 1, 2, \dots, n)$$

besitzt den Rang $n-l$.

Aus den voraufgehenden Theoremen ergeben sich folgende, besonders häufig benützte Sätze:

Damit ein System von n linearen homogenen Gleichungen mit der gleichen Anzahl von Unbekannten Lösungen besitzt, die nicht sämtlich Null sind, ist das Verschwinden der aus den Koeffizienten des Gleichungssystemes gebildeten Determinante notwendig und hinreichend. Ein System linearer homogener Gleichungen mit weniger Gleichungen als Unbekannten kann stets durch Werte befriedigt werden, die nicht sämtlich Null sind.

Hat man n — 1 lineare homogene Gleichungen:

$$a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n = 0 \qquad (i = 1, 2, \dots, n-1)$$

mit n Unbekannten und hat die zugehörige Matrix der Koeffizienten

$$\| a_{ik} \| \qquad (i = 1, 2, \ldots, n-1;\ k = 1, 2, \ldots, n)$$

den Rang $n - 1$, so ergibt sich:

$$x_1 : x_2 : \cdots : x_n = A_1 : A_2 : \cdots : A_n;$$

$A_1, A_2, \ldots, A_n$ sind die mit abwechselnden Vorzeichen genommenen Determinanten, die aus der Matrix durch sukzessives Streichen der ersten, zweiten, usw. letzten Kolonne hervorgehen.

Bei der Behandlung eines Systemes linearer Gleichungen gelangte Leibniz (*Brief an l'Hospital,* 1693) zu den Determinanten. Die Auflösung linearer Gleichungen mittels Determinanten in dem sogenannten allgemeinen Fall, d. h. dem, der nicht alle Besonderheiten umfaßt, ist von Cramer gegeben worden. Für ein System von n linearen Gleichungen mit der nämlichen Anzahl von Unbekannten sagt noch Jacobi (*de formatione et proprietatibus determinantium,* Art. 7 (1841), deutsche Ausg. von Stäckel in *Ostwalds Klass. der exakten Wiss.*): „Man hat also, wenn die Determinante verschwindet, noch eine Mannigfaltigkeit von Fällen sehr verschiedener Natur zu unterscheiden, und man müßte algebraische Kriterien für die einzelnen Fälle angeben. Das scheint jedoch für eine beliebige Anzahl linearer Gleichungen recht weitläufig zu sein.“ Die allgemeine Behandlung linearer homogener Gleichungen mit Hülfe von Unterdeterminanten hat zuerst Kronecker gegeben (Baltzer, *Determinanten,* 2. Aufl. (1864), S. 62); man vgl. ferner Frobenius, *Journ. f. Math.* **82**, 236 (1877). Die Bedingungen für die Auflösbarkeit eines Systemes linearer unhomogener Gleichungen geben in allgemeiner Form Fontené, *Nouv. ann.* (2) **14** (1875), sowie Rouché, *C. R.* **81** (1875). Der Begriff „Rang“ stammt von Frobenius, *Journ. f. Math.* **82**, 239 (1877) u. **86**, 148 (1879), vgl. auch ebenda **129**, 175 (1905).

Von Lehrbüchern verweisen wir auf die Darstellungen bei H. Weber, *Algebra* **1**, 96, Kronecker, *19. Vorl. über Determ.* Gordan, *Vorl. über Invariantentheorie* **1**, 101, E. v. Weber *Vorl. über das Pfaffsche Problem, Teubners Sammlung* 1900, S. 6.

§ 6. Das Rechnen mit Matrices. Zusammenhang zwischen Matrices, linearen Substitutionen und bilinearen Formen.

Die Gesamtheit der n^2 in einem quadratischen Schema

$$\begin{Vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{Vmatrix}$$

angeordneten Elemente a_{ik}, bei denen der erste Index die Zeile, der zweite die Kolonne charakterisiert, heißt eine *quadratische Matrix vom Grade oder der Ordnung n* und *soll nach Cayley zusammengefaßt mit einem einzigen Buchstaben A bezeichnet werden. „Matrices gleichen Grades verhalten sich, wie man sehen wird, wie einzelne Größen"* (Cayley, *Phil. Trans.* (1858), *Coll. math. papers* 2, 475). Die im folgenden betrachteten Matrices sollen *stets von gleichem Grade* vorausgesetzt werden. Sollte dies von Anfang an nicht der Fall sein, so füge man den Matrices niedrigeren Grades, um den Defekt zu beseitigen, Zeilen und Kolonnen bei, deren Elemente lauter Nullen sind. *Auf die nämliche Weise sollen rechteckige Matrices in quadratische umgestaltet werden, so daß die Voraussetzung, die Matrices sollen quadratisch sein, keine Beschränkung involviert.*

Für quadratische Matrices gleichen Grades kann man ein *symbolisches Rechnen* definieren.

Eine Gleichung $A = B$ zwischen zwei Matrices n^{ten} Grades soll besagen, daß jedes der n^2 Elemente von A gleich dem entsprechenden von B ist. Sind die Elemente von B mit b_{ik} bezeichnet, so ist die Gleichung $A = B$ mit den n^2 Gleichungen

$$a_{ik} = b_{ik} \qquad (i, k = 1, 2, \ldots, n)$$

gleichbedeutend.

Die *Addition* zweier beliebiger Matrices A und B gleichen Grades n mit den Elementen a_{ik} und b_{ik} wird auf folgende Weise definiert: Man bilde die neue Matrix S vom n^{ten} Grad mit den n^2 Elementen $s_{ik} = a_{ik} + b_{ik}$. *Die Matrix S heißt die Summe von A und B; man schreibt $S = A + B$.* Diese Summenbildung ist *kommutativ*, d. h. es besteht die Gleichung

$$A + B = B + A.$$

Für drei Matrices A, B, C gilt das *assoziative* Gesetz, d. h. es ist

$$A + (B + C) = (A + B) + C.$$

Aus zwei Matrices A mit den Elementen a_{ik} und B mit den Elementen b_{ik} von gleichem Grade n kann eine neue Matrix Q mit den Elementen q_{ik} von dem nämlichen Grade gebildet werden, deren Elemente q_{ik} gleich

$$a_{i1}b_{1k} + a_{i2}b_{2k} + \cdots + a_{in}b_{nk} \qquad (i, k = 1, 2, \ldots, n)$$

definiert sind. *Diese Matrix Q heißt aus A und B komponiert;* man sagt *Q ist das Produkt von A und B und schreibt symbolisch $Q = AB$.*

Die symbolische Gleichung $Q = AB$ ist gleichbedeutend mit den n^2 Gleichungen

$$q_{ik} = \sum_{t=1}^{t=n} a_{it}b_{tk} \qquad (i, k = 1, 2, \ldots, n);$$

die Elemente von Q entstehen aus denjenigen der Matrices A und B, indem man die Zeilen von A mit den Kolonnen von B komponiert. In den Anwendungen ist nur diese Art der Verknüpfung von Wichtigkeit. *Bedeutet $|Q|$ die Determinante von Q, so ist nach dem Multiplikationssatz für Determinanten:*

$$|Q| = |AB| = |A| \cdot |B|.$$

Ferner ist als Zeilenprodukt zweier rechteckiger Matrices (zu vgl. S. 60) die Determinante

$$\sum \pm q_{g_1 h_1} q_{g_2 h_2} \cdots q_{g_m h_m} = \sum_{\sigma_1 \sigma_2 \ldots \sigma_m} A^{g_1 g_2 \ldots g_m}_{\sigma_1 \sigma_2 \ldots \sigma_m} \cdot B^{\sigma_1 \sigma_2 \ldots \sigma_m}_{h_1 h_2 \ldots h_m};$$

in der Summe auf der rechten Seite ist jede Kombination der Zahlen 1, 2, ..., n zu m zu setzen.

$A^{g_1 g_2 \ldots g_m}_{\sigma_1 \sigma_2 \ldots \sigma_m}$ bedeutet die Determinante $\sum \pm a_{g_1 \sigma_1} a_{g_2 \sigma_2} \ldots a_{g_m \sigma_m}$,

$B^{\sigma_1 \sigma_2 \ldots \sigma_m}_{h_1 h_2 \ldots h_m}$ bedeutet die Determinante $\sum \pm b_{\sigma_1 h_1} b_{\sigma_2 h_2} \ldots b_{\sigma_m h_m}$.

Jede Unterdeterminante m^{ten} Grades von $|AB|$ ist demnach eine ganze lineare homogene Funktion sowohl der Unterdeterminanten m^{ten} Grades von $|A|$ als auch der von $|B|$.

Die definierte symbolische Multiplikation erfüllt das *assoziative Gesetz*, d. h. sind A, B, C drei Matrices gleichen Grades, so ist $A(BC) = (AB)C$.

Die definierte Multiplikation genügt auch dem *distributiven* Gesetz, d. h. für irgend drei Matrices A, B, C gleichen Grades gelten die Relationen:

$$(A + B)C = AC + BC,$$
$$C(A + B) = CA + CB.$$

Hingegen ist die erklärte symbolische Multiplikation im allgemeinen *nicht kommutativ*, d. h. AB ist im allgemeinen von BA verschieden. Findet im besonderen für zwei Matrices A und B die Gleichung $AB = BA$ statt, so heißen A und B *vertauschbar* oder *kommutativ*.

Eine Matrix, bei der alle Elemente mit Ausnahme der Diagonalelemente Null sind, diese aber sämtlich den nämlichen Wert ϱ haben, heißt eine *Diagonalmatrix* und soll durch die in der Diagonale enthaltene Zahl ϱ bezeichnet werden. Diejenige Diagonalmatrix, bei der alle Diagonalelemente gleich 1 sind, heißt die *Einheitsmatrix* und soll nach Frobenius' grundlegender Arbeit (*Über lineare Substitutionen und bilineare Formen, Journ. f. Math.* 84, 1 (1878)) durchgehend mit E bezeichnet werden. Ist A eine beliebige Matrix und ϱ eine Diagonalmatrix, so ist $\varrho A = A\varrho$, d. h. *eine Diagonalmatrix ist mit jeder Matrix vertauschbar;* $A\varrho$ hat die Elemente $\varrho \cdot a_{ik}$. Im besonderen ist für die Einheitsmatrix: $AE = EA = A$.

Nur wenn die Determinante einer Matrix A nicht verschwindet, kann eine Matrix Z der Gleichung $AZ = E$ genügen. *Für jede Matrix A von nicht verschwindender Determinante existiert eine eindeutig bestimmte Matrix Z, für die $AZ = E$ wird; sie ist mit A vertauschbar und soll mit A^{-1} bezeichnet werden. Diese Matrix A^{-1}, deren Elemente*

$$a_{ik} = \frac{A_{ki}}{|A|}$$

lauten, heißt die zur Matrix A reziproke oder inverse Matrix. A_{ik} ist hierbei die algebraische Adjungierte des Elementes a_{ik} der Determinante $|A|$, also

$$A_{ik} = \frac{\partial |A|}{\partial a_{ik}}.$$

Die reziproke Matrix zu A^{-1} ist A, d. h. $(A^{-1})^{-1} = A$, die reziproke Matrix des Produktes AB lautet $B^{-1}A^{-1}$.

Unter A^r versteht man die Matrix, die man durch r malige Zusammensetzung von A mit sich selbst erhält; *sie wird die*

r^{te} Potenz von A genannt. Ist die Determinante von A nicht Null, so entsteht durch r malige Zusammensetzung von A^{-1} eine Matrix, die mit A^{-r} bezeichnet wird. A^r und A^{-r} sind reziprok. Sowohl für positive als auch für negative Werte der Exponenten r_1 und r_2 gilt die Relation:

$$A^{r_1} \cdot A^{r_2} = A^{r_2} \cdot A^{r_1} = A^{r_1 + r_2}.$$

Zu jeder quadratischen Matrix A gehört eine *lineare homogene Substitution. Eine lineare homogene Substitution n^{ten} Grades ist diejenige Operation, die n Variable durch n lineare homogene Funktionen von n neuen Variablen ersetzt.* Bezeichnet man die n ursprünglichen Variablen mit $y_1, y_2, \ldots, y_n$, die n neuen mit $x_1, x_2, \ldots, x_n$, so lautet die lineare homogene Substitution, deren Koeffizienten man sich durch die Matrix A gegeben denkt:

$$\begin{aligned} y_1 &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n, \\ y_2 &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n, \\ &\vdots \\ y_n &= a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n. \end{aligned}$$

Bei einer Substitution ist die Variablenbezeichnung unwesentlich; die Eigenschaften werden ausschließlich durch das System der Koeffizienten a_{ik} bedingt. Daher bezeichnet man die Substitution einfach mit demselben Buchstaben A wie die zugehörige Matrix. Sollen bei abgekürzter Bezeichnung der Substitution auch die Variablen hervorgehoben werden, so schreibt man:

$$(y) = A\,(x) \text{ oder } (y_1, y_2, \ldots, y_n) = A\,(x_1, x_2, \ldots, x_n).$$

Die Gleichung $(y) = A\,(x)$ ist im Sinne der Gleichheit von Matrices erfüllt, wenn man unter (x) die Matrix

$$\left\| \begin{matrix} x_1 & 0 & 0 & \ldots & 0 \\ x_2 & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ x_n & 0 & 0 & \ldots & 0 \end{matrix} \right\|$$

und unter (y) die analoge Matrix versteht und $A\,(x)$ als Produkt der zwei Matrices A und (x) auffaßt (vgl. Molien, *Math. Ann.* **41**, 149 (1893) oder Wellstein, *Arch. f. Math.* (3) **5**, 230 (1903)).

Die Anzahl der linear unabhängigen unter den Größen $y_1, y_2, \ldots, y_n$ ist gleich dem Range l der Matrix A (zu vgl. S. 63). Hierzu ist zu bemerken, daß man von $l + 1$ $(l < n)$

linearen homogenen Funktionen $y_{\alpha_1}, y_{\alpha_2}, \ldots, y_{\alpha_{l+1}}$ der n Variablen $x_1, x_2, \ldots, x_n$ sagt, sie sind *in linearer Dependenz,* falls es Konstante $c_{\alpha_1}, c_{\alpha_2}, \ldots, c_{\alpha_{l+1}}$ gibt, die nicht sämtlich Null sind, so daß identisch, d. h. für jeden Wert von $x_1, x_2, \ldots, x_n$:

$$c_{\alpha_1} y_{\alpha_1} + c_{\alpha_2} y_{\alpha_2} + c_{\alpha_3} y_{\alpha_3} + \cdots + c_{\alpha_{l+1}} y_{\alpha_{l+1}} = 0$$

wird; die lineare Abhängigkeit von $y_{\alpha_1}, y_{\alpha_2}, \ldots, y_{\alpha_{l+1}}$ erfordert also, daß die n linearen homogenen Gleichungen:

$$c_{\alpha_1} a_{\alpha_1 i} + c_{\alpha_2} a_{\alpha_2 i} + \cdots + c_{\alpha_{l+1}} a_{\alpha_{l+1}, i} = 0 \quad (i = 1, 2, \ldots, n)$$

ein Lösungssystem $c_{\alpha_1}, c_{\alpha_2}, \ldots, c_{\alpha_{l+1}}$ haben, dessen Elemente c nicht sämtlich Null sind. Ist im besonderen die Determinante

$$\sum \pm a_{11}\, a_{22} \cdots a_{ll}$$

nicht Null und die Determinante von A:

$$|A| = \sum \pm a_{11}\, a_{22} \cdots a_{nn}$$

vom Range l, so sind die ersten l Funktionen $y_1, y_2, \ldots, y_l$ linear unabhängig und $y_{l+1}, y_{l+2}, \ldots, y_n$ lassen sich durch $y_1, y_2, \ldots, y_l$ linear und homogen mit konstanten Koeffizienten ausdrücken.

Die Determinante

$$|A| = \sum \pm a_{11}\, a_{22} \cdots a_{nn}$$

der Matrix A heißt die *Substitutionsdeterminante* oder der *Modul der Substitution. Ist die Substitutionsdeterminante* $|A| \neq 0$, *so kann man die Substitution* $(y) = A(x)$ *nach den Variablen* $x_1, x_2, \ldots, x_n$ *auflösen und findet:* $(x) = A^{-1}(y)$; *hierbei ist* A^{-1} *die zu* A *reziproke Matrix.* Die zwei Substitutionen $(y) = A(x)$ und $(x) = A^{-1}(y)$ heißen *reziprok* oder *invers.*

Hat man eine zu der Matrix A zugehörige Substitution $(x) = A(x')$ und führt für die Variablen $x_1', x_2', \ldots, x_n'$ vermöge einer zur Matrix B gehörigen Substitution $(x') = B(x'')$ neue Variablen $x_1'', x_2'', \ldots, x_n''$ ein, so erhält man eine neue Substitution $(x) = Q(x'')$, deren Koeffizienten

$$q_{ik} = a_{i1} b_{1k} + a_{i2} b_{2k} + \cdots + a_{in} b_{nk}$$

lauten. *Die resultierende Substitution* Q *besitzt mithin die Matrix* AB *und wird daher auch als Produkt der zwei Substitutionen* A *und* B *bezeichnet.*

Hat man eine lineare homogene Substitution $(x) = A(x')$, so ist es oft zweckmäßig, für die Variablen x und x' neue Variablen y und y' durch zwei lineare *kogrediente*[1]) Substitutionen, d. h. solche mit gleichen Matrices, einzuführen:

$$y_i = p_{i1}x_1 + p_{i2}x_2 + \cdots + p_{in}x_n \qquad (i = 1, 2, \ldots, n)$$
$$y_i' = p_{i1}x_1' + p_{i2}x_2' + \cdots + p_{in}x_n' \qquad (i = 1, 2, \ldots, n)$$

oder abgekürzt:

$$(y) = P(x),$$
$$(y') = P(x'),$$

und die resultierende Substitution $(y) = C(y')$ zu betrachten. *Von der Matrix P wird hierbei nur vorausgesetzt, daß die Determinante von P nicht verschwindet.* Infolge der Voraussetzung $|P| \neq 0$ existiert $(x) = P^{-1}(y)$, und man findet: $C = PAP^{-1}$.

Hat man irgend zwei Matrices oder Substitutionen A und C und kann man eine Matrix oder Substitution P von nicht verschwindender Determinante finden, daß $C = PAP^{-1}$ wird, so heißt C mit A ähnlich, konjugiert oder gleichberechtigt. Ähnliche Substitutionen werden bisweilen auch als *äquivalent* bezeichnet; jedoch wird der Begriff „äquivalent" auch im weiteren Sinne (vgl. S. 89) verwandt und soll daher im folgenden *nicht* als mit ähnlich gleichbedeutend benützt werden.

Aus $C = PAP^{-1}$ folgt $A = P^{-1}CP$ und, da $(P^{-1})^{-1} = P$ ist, so ergibt sich, daß *die Ähnlichkeit von A und C eine gegenseitige ist.*

Alle mit einer Substitution ähnlichen Substitutionen bilden eine Klasse ähnlicher Substitutionen, so daß zwei Substitutionen derselben Klasse stets untereinander ähnlich sind.

Bezeichnet ϱ eine Diagonalmatrix n^{ten} Grades und sind A und C zwei ähnliche Systeme n^{ten} Grades, so sind auch $\varrho E - A$ und $\varrho E - C$ ähnliche Matrices, wie aus

$$\varrho E - C = \varrho E - PAP^{-1} = P(\varrho E - A)P^{-1}$$

folgt. Die nähere Untersuchung der zu einer Substitution A ähnlichen Substitutionen verlangt daher die Betrachtung der Determinante der Matrix $\varrho E - A$, d. h. der Determinante mit den Elementen

$$\varrho\delta_{ik} - a_{ik} \qquad (\delta_{ii} = 1, \delta_{ik} = 0 \text{ für } i \gtrless k).$$

1) Anstatt kogredient verwendet man auch die Bezeichnung „*kongruent*" (Kronecker, *Berl. Monatsb.* 1874, *Ges. Werke* **1**, 424).

Für irgendeine Substitution A heißt *die Determinante* $|\varrho E - A|$, *die eine ganze rationale Funktion n^{ten} Grades von ϱ ist*, in Anlehnung an Cauchy *die charakteristische Funktion der Substitution oder Matrix A* (Frobenius, *Journ. f. Math.* **84**, 10); entsprechend heißt die Gleichung $|\varrho E - A| = 0$ die *charakteristische Gleichung* oder nach L. Fuchs (ebenda **66**, 133 (1866), *Ges. Werke* **1**, 172) die *Fundamentalgleichung der Substitution A. Grenzen für die reellen und imaginären Teile der Wurzeln der Fundamentalgleichung* bei J. Bendixson u. A. Hirsch, *Acta math.* **25**, 359 u. 367 (1902) u. Bromwich, ebenda, **30**, 297 (1906).

Ist $\varphi(\varrho) = |\varrho E - A|$ die charakteristische Funktion der Matrix A, die vom n^{ten} Grade sei, und $\vartheta(\varrho)$ der größte gemeinsame Teiler aller Unterdeterminanten $n-1^{ten}$ Grades der Determinante $|\varrho E - A|$, so ist

$$\lambda(\varrho) = \frac{\varphi(\varrho)}{\vartheta(\varrho)}$$

eine ganze rationale Funktion von ϱ; sie heißt der n^{te} Elementarteiler der charakteristischen Funktion von A (vgl. S. 103). Sei $\lambda(\varrho) = \lambda_0 \varrho^q + \lambda_1 \varrho^{q-1} + \cdots + \lambda_q$, so besteht die Gleichung:

$$\lambda_0 A^q + \lambda_1 A^{q-1} + \cdots + \lambda_q = 0.$$

Diese Gleichung ist *die Gleichung niedrigsten Grades, der die Matrix A genügt;* sie heißt die *reduzierte charakteristische Gleichung der Matrix A* oder auch *die Grundgleichung.* Ist $\chi(A) = 0$ irgendeine Gleichung, der die Matrix A genügt, so ist $\chi(\varrho)$ durch $\lambda(\varrho)$ teilbar. Dieser grundlegende Satz der Theorie der Matrices stammt von Frobenius, *Journ. f. Math.* **84**, 11, vgl. besonders *Sitzungsb. d. Berl. Akad.* (1896), 606, ferner Ed. Weyr, *Monatsh. f. Math.* **1**, 187 (1890), sowie O. Perron, *Math. Ann.* **64**, 249 (1907). Da $\varphi(\varrho) = \lambda(\varrho) \cdot \vartheta(\varrho)$ und $\lambda(A) = 0$ ist, so ist natürlich $\varphi(A) = 0$. Die Eigenschaft jeder Matrix A, ihrer charakteristischen Gleichung

$$\varphi(\varrho) = \varrho^n + \varphi_1 \varrho^{n-1} + \cdots + \varphi_n = 0 \quad (\varphi_n = (-1)^n |A|)$$

zu genügen, ist zuerst von Cayley in der auf S. 79 genannten Arbeit (*Coll. math. papers* **2**, 482) ausgesprochen worden.

In innigstem Zusammenhang mit der charakteristischen Gleichung steht der Begriff eines *linear unabhängigen Systemes von Matrices.* Ein System von Matrices $A_1, A_2, \ldots, A_m$ heißt linear abhängig, wenn m numerische Konstanten $\mu_1, \mu_2, \ldots, \mu_m$

existieren, die nicht gleichzeitig verschwinden, so daß

$$\mu_1 A_1 + \mu_2 A_2 + \cdots + \mu_m A_m = 0$$

wird. Ist allgemein:

$$A_\alpha = \| a_{ik}^{(\alpha)} \| \quad (\alpha = 1, 2, \ldots, m;\ i, k = 1, 2, \ldots, n),$$

so ist die obige symbolische Gleichung gleichbedeutend mit den n^2 gewöhnlichen Gleichungen:

$$\mu_1 a_{ik}^{(1} + \mu_2 a_{ik}^{(2)} + \cdots + \mu_m a_{ik}^{(m)} = 0 \quad (i, k = 1, 2, \ldots, n).$$

Jedes nicht linear abhängige System von Matrices oder bilinearen Formen heißt *linear unabhängig*. Mehr als n^2 Matrices n^{ten} Grades sind stets linear abhängig. Unter den Matrices n^{ten} Grades lassen sich auf unendlich viele Weisen n^2 linear unabhängige angeben; ein solches System $A_1, A_2, \ldots, A_{n^2}$ hat die Eigenschaft, daß jede Matrix C in der Form

$$C = c_1 A_1 + c_2 A_2 + \cdots + c_{n^2} A_{n^2}$$

erscheint, wobei $c_1, c_2, \ldots, c_{n^2}$ numerische Konstanten bedeuten. Die Existenz der charakteristischen Gleichung besagt, daß unter den Potenzen $A^0, A^1, A^2, \ldots, A^{n-1}, A^n$ einer Matrix n^{ten} Grades höchstens n linear unabhängig sind.

Verschwindet die charakteristische Gleichung einer Matrix A für die s verschiedenen Werte $a_1, a_2, \ldots, a_s$, so kann man s Matrices, die *„Frobeniusschen Kovarianten"*, konstruieren. Diese Matrices besitzen die Eigenschaft, daß keine von ihnen verschwindet, sie ferner linear unabhängig sind,

$$E = K_1 + K_2 + \cdots + K_s,\ K_i^2 = K_i,\ K_i K_k = 0 \quad (i \gtrless k)$$

und

$$A = a_1 K_1 + a_2 K_2 + \cdots + a_s K_s + A_0$$

wird. Hat die reduzierte charakteristische Gleichung lauter verschiedene Wurzeln, so ist $A_0 = 0$, sonst wird wenigstens eine Potenz von A_0 identisch Null. Vgl. Frobenius, *Über die schiefe Invariante einer bilin. od. quadr. Form, Journ. f. Math.* **86**, 44 (1879), *Über vertauschbare Matrices, Sitzungsb. d. Berl. Akad.* (1896), 609, Study, *Rekurrierende Reihen und bilin. Form., Monatsh. f. Math.* **2**, 23 (1891), Wellstein, *Über die Frobeniusschen Kovarianten, Arch. f. Math.* (3) **5**, 229 (1903).

Über die Elementarteiler der charakteristischen Funktion und die Normalform linearer homogener Substitutionen vgl. § 8.

Vertauscht man die Zeilen und Kolonnen der Matrix A, so erhält man die zu A *transponierte* oder *konjugierte Matrix oder Substitution.* Sie wird mit A' bezeichnet; ihre Koeffizienten a'_{ik} sind gleich a_{ki}. *Die zu A' transponierte Substitution ist die ursprüngliche Substitution A. Sind A' und B' die zu A und B transponierten Substitutionen, so ist $B'A'$ die transponierte Substitution von AB. Der Zusammenhang zwischen reziproker und transponierter Substitution drückt sich durch die Gleichung $(A^{-1})' = (A')^{-1}$ aus.*

Sind

$$\alpha_{ik} = \frac{A_{ki}}{|A|}$$

die Koeffizienten der zu A reziproken Substitution A^{-1}, so hat A'^{-1} in der i^{ten} Zeile und k^{ten} Kolonne die Koeffizienten

$$\frac{A_{ik}}{|A|}.$$

Die Substitution A'^{-1} heißt zu A *kontragredient.*

Notwendig und hinreichend, damit zwei Substitutionen

$$(y) = A(x) \quad \text{und} \quad (Y) = B(X)$$

kontragredient sind, ist das Bestehen der Identität

$$y_1 Y_1 + y_2 Y_2 + \cdots + y_n Y_n = x_1 X_1 + x_2 X_2 + \cdots x_n X_n.$$

Sind A'^{-1} und B'^{-1} die zu A und B kontragredienten Substitutionen, so lautet die zu AB kontragrediente $A'^{-1}B'^{-1}$.

Unter $\bar{A}$ versteht man diejenige Substitution, die aus A hervorgeht, indem man alle Koeffizienten von A durch die zu ihnen konjugiert imaginären ersetzt. $\bar{A}$ heißt die zu A konjugiert imaginäre Substitution.

Hat A im besonderen reelle Koeffizienten, so ist $\bar{A}$ mit A identisch.

Jeder Matrix A vom n^{ten} Grade mit den Elementen a_{ik} läßt sich eine *bilineare Form*

$$\sum_{i=1}^{i=n} \sum_{k=1}^{k=n} a_{ik} x_i y_k$$

zuordnen (das Studium solcher Formen beginnt mit Jacobis Aufsatz im *Journ. f. Math.* **53**, 265 (1857), *Ges. Werke* **3**, 583). *Eine solche Form hängt von zwei Reihen von Variablen $x_1, x_2, \ldots, x_n$ und $y_1, y_2, \ldots, y_n$ ab.* Ebenso wie die aus den Elementen a_{ik}

gebildete quadratische Matrix oder eine dem Koeffizientensystem a_{ik} zugeordnete lineare Substitution soll auch die bilineare Form

$$\sum_{i=1}^{i=n}\sum_{k=1}^{k=n} a_{ik}x_iy_k$$

mit dem Buchstaben A bezeichnet werden. (Die Verknüpfung von Matrices und bilinearen Formen verdankt man Frobenius, *Journ. f. Math.* **84**, 1.)

Hat man irgendeine bilineare Form

$$A = \sum_{i=1}^{i=n}\sum_{k=1}^{k=n} a_{ik}x_iy_k,$$

so heißt

$$A' = \sum_{i=1}^{i=n}\sum_{k=1}^{k=n} a_{ki}x_iy_k$$

die zu A *transponierte* oder *konjugierte* bilineare Form. Wird unter $\bar{a}_{ik}$ stets die zu a_{ik} konjugiert imaginäre Größe verstanden, so heißt

$$\bar{A} = \sum_{i=1}^{i=n}\sum_{k=1}^{k=n} \bar{a}_{ik}x_iy_k$$

die zu A *konjugiert imaginäre bilineare Form.* Ist die Determinante $|A|$ von A von Null verschieden, so existiert die zu A *reziproke* oder *inverse* Form

$$A^{-1} = \sum_{i=1}^{i=n}\sum_{k=1}^{k=n} \frac{A_{ki}}{|A|}x_iy_k$$

sowie die zu A *kontragrediente bilineare Form*

$$A'^{-1} = \sum_{i=1}^{i=n}\sum_{k=1}^{k=n} \frac{A_{ik}}{|A|}x_iy_k.$$

Eine bilineare Form

$$A = \sum_{i=1}^{i=n}\sum_{k=1}^{k=n} a_{ik}x_iy_k$$

heißt *symmetrisch*, falls $a_{ik} = a_{ki}$, und *alternierend*, falls $a_{ik} = -a_{ki}$ $(i \gtreqless k)$, also im besonderen $a_{ii} = 0$. Die Form A wird durch die symbolische Gleichung $A = A'$ als symme-

trisch, durch $A = -A'$ als alternierend charakterisiert. Die symbolische Gleichung $A' = \bar{A}$ definiert eine *Hermitesche Form.*

Transformiert man eine bilineare Form

$$D = \sum_{i=1}^{i=n} \sum_{k=1}^{k=n} d_{ik} x_i y_k$$

(die Bezeichnung ist mit Rücksicht auf § 8 gewählt, zu vgl. S. 106) durch die zwei linearen homogenen Substitutionen:

$$P: x_i = p_{i1} x_1' + p_{i2} x_2' + \cdots + p_{in} x_n' \qquad (i = 1, 2, \ldots, n)$$

und

$$Q: y_i = q_{i1} y_1' + q_{i2} y_2' + \cdots + q_{in} y_n' \qquad (i = 1, 2, \ldots, n)$$

in die bilineare Form

$$F = \sum_{i=1}^{i=n} \sum_{k=1}^{k=n} f_{ik} x_i' y_k',$$

so sagt man, *die Form D geht durch die Substitutionen P und Q in F über. Dann bestehen die n^2 Gleichungen:*

$$f_{ik} = \sum_{s=1}^{s=n} \sum_{t=1}^{t=n} p_{si} d_{st} q_{tk} \qquad (i, k = 1, 2, \ldots, n)$$

oder

$$f_{ik} = \sum_{s=1}^{s=n} \sum_{t=1}^{t=n} p'_{is} d_{st} q_{tk} \qquad (i, k = 1, 2, \ldots, n),$$

wenn man $p'_{ik} = p_{ki}$ setzt. Ist P' das zu P transponierte System, so kann man die n^2 Gleichungen in die eine symbolische Gleichung $F = P'DQ$ zusammenfassen.

Besteht zwischen zwei Matrices F und D eine Gleichung $F = RDQ$, d. h. kann man F dadurch aus D herleiten, daß man D vorn und hinten mit zwei beliebigen Matrices R und Q komponiert, so heißt F ein *Vielfaches* von D und D *ein Teiler von F. Zwei Systeme D und F heißen äquivalent, wenn sowohl F ein Vielfaches von D, als auch D ein Vielfaches von F ist.*

Bei den Begriffen des Vielfachen und der Äquivalenz führt man für die Koeffizienten von R und Q auch noch Beschränkungen ein, die von der Natur der Koeffizienten von D und F abhängen, wie z. B. die Koeffizienten von R und Q sollen reell oder ganzzahlig sein. In diesem Paragraphen sollen die Koeffizienten von R und Q keiner Beschränkung unterworfen werden.

sondern beliebige Größen sein können. Der § 8 wird von *besonderen Arten der Äquivalenz* handeln.

Der Rang des Vielfachen F ist stets gleich oder kleiner als der Rang des Teilers D.

Zwei Matrices F und D sind dann und nur dann äquivalent, wenn sie gleichen Rang haben; in diesem Fall gibt es zwei Matrices R und Q von nichtverschwindenden Determinanten, so daß einerseits $F = RDQ$ und andererseits wegen des Nichtverschwindens der Determinanten $|R|$ und $|Q|$ auch $D = R^{-1}FQ^{-1}$ wird. Im besonderen ist eine Matrix A vom Range l der Einheitsmatrix l^{ten} Grades mit den Elementen

$$\delta_{ii} = 1, \ \delta_{ik} = 0 \qquad (i, k = 1, 2, \ldots, l)$$

äquivalent.

Setzt man $P' = R$, so hat man die Ausgangsgleichung $F = P'DQ$. Überträgt man die letzten Sätze in die Sprache der bilinearen Formen, so erhält man folgende Theoreme:

Hat man eine bilineare Form

$$D = \sum_{i=1}^{i=n} \sum_{k=1}^{k=m} d_{ik} x_i y_k$$

(durch Hinzunahme von Elementen 0 kann man zum Zweck der symbolischen Produktbildung die bilineare Form mit zwei Reihen von gleichvielen Variablen geschrieben denken) *vom Range l und transformiert sie durch die zwei linearen homogenen Substitutionen:*

$$P: x_i = p_{i1} x_1' + p_{i2} x_2' + \cdots + p_{in} x_n' \qquad (i = 1, 2, \ldots, n)$$

und

$$Q: y_i = q_{i1} y_1' + q_{i2} y_2' + \cdots + q_{im} y_m' \qquad (i = 1, 2, \ldots, m)$$

in die neue bilineare Form

$$F = \sum_{i=1}^{i=n} \sum_{k=1}^{k=m} f_{ik} x_i' y_k',$$

so hat diese den gleichen oder niedrigeren Rang wie D. Verschwinden die Determinanten von P und Q nicht, so haben, da man dann auch F durch P^{-1} und Q^{-1} in D zurücktransformieren kann, D und F den gleichen Rang. Man kann im besonderen stets zwei derartige Substitutionen P und Q von nicht verschwindenden Determinanten auswählen, daß D in die Einheitsform $x_1' y_1' + x_2' y_2' + \cdots + x_l' y_l'$ übergeht.

Zwei bilineare Formen gleichen Ranges lassen sich stets durch zwei lineare homogene Transformationen von nicht verschwindenden Determinanten ineinander überführen.

Damit zwei bilineare Formen, die von gleichvielen Variablenpaaren abhängen, durch *kogrediente* Transformationen, d. h. durch die nämlichen Substitutionen für die zwei Variablenreihen, ineinander überführbar sind (also $P = Q$), müssen die Formen D und F, außer im Range übereinzustimmen, noch *weiteren* Bedingungen genügen (zu vgl. S. 116). *Für zwei symmetrische oder alternierende bilineare Formen ist die bloße Übereinstimmung des Ranges ihrer Determinanten für die kogrediente Transformation ausreichend.*

Hat man eine rechteckige oder quadratische Matrix

$$A = \| a_{ik} \| \qquad (i = 1, 2, \ldots, n;\ k = 1, 2, \ldots, m),$$

so ist es bisweilen vorteilhaft, sie in der Form:

$$\left\| \begin{matrix} A_{11} & A_{12} & \ldots & A_{1p} \\ A_{21} & A_{22} & \ldots & A_{2p} \\ \vdots & \vdots & & \vdots \\ A_{q1} & A_{q2} & \ldots & A_{qp} \end{matrix} \right\|$$

zu schreiben. Die A mit zwei unteren Indices stellen selbst *rechteckige Matrices* dar, alle A mit dem gleichen ersten Index haben dieselbe Anzahl Horizontalreihen, alle A mit dem nämlichen zweiten Index besitzen die gleiche Anzahl Vertikalreihen. In den Anwendungen ist der Fall am wichtigsten, daß $A_{11}, A_{22}, \ldots, A_{qp}$ lauter quadratische Matrices sind. Sind alle Elemente einer Matrix A_{ik} Null, so setzt man für A_{ik} einfach 0.

Ist A eine Matrix, die sich auf Diagonalsysteme reduziert und deren andere Elemente ausschließlich Nullen sind:

$$A = \left\| \begin{matrix} A_{11} & 0 & 0 & \ldots & 0 \\ 0 & A_{22} & 0 & \ldots & 0 \\ 0 & 0 & A_{33} & \ddots & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & & A_{pp} \end{matrix} \right\|,$$

so verwendet man die von A. Hurwitz (H. Kreis, *Contribution à la théorie des systèmes linéaires, Thèse,* Zürich 1906) angegebene Bezeichnung

$$A_{11} \dotplus A_{22} \dotplus \cdots \dotplus A_{pp},$$

oder schreibt auch

$$\{A_{11}, A_{22}, \ldots, A_{pp}\}.$$

Analog wie bilineare Formen mit zwei Reihen von verschieden vielen Variablen (S. 90) kann man auch eine lineare homogene Substitution mit zwei Reihen von verschieden vielen Variablen einführen:

$$\begin{aligned} y_1 &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m, \\ y_2 &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2m}x_m, \\ &\vdots \qquad\quad \vdots \qquad\quad\quad \vdots \\ y_n &= a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nm}x_m; \end{aligned}$$

man kann ihr die rechteckige Matrix

$$A_{11} = \left\| \begin{matrix} a_{11} & a_{12} & \cdots & a_{1m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{matrix} \right\|$$

oder zum Zweck der symbolischen Rechnung für $n > m$ die quadratische Matrix

$$\| A_{11}\, 0 \|,$$

für $n < m$ die quadratische Matrix

$$\left\| \begin{matrix} A_{11} \\ 0 \end{matrix} \right\|$$

zuordnen.

Das symbolische Rechnen mit Matrices ist eine Schöpfung Cayleys (vgl. den Anfang dieses Paragraphen); inwieweit W. R. Hamilton durch seine *Lectures on quaternions* (1853) als Cayleys Vorgänger anzusehen ist, entnehme man den historischen Bemerkungen von Taber (*Am. J. math.* **12**, 337 (1890)) und von Bromwich (*Bull. Am. math. Soc.* **7**, 311 (1901)), sowie dem Schluß des § 7. Laguerres *„Calcul des systèmes linéaires"* (*J. éc. polyt.*, Cah. **42**, 215 (1867), *Œuvres* **1**, 221) ist Cayleys Untersuchungen analog. Die tiefgehendste Behandlung des in diesem Paragraphen besprochenen Gegenstandes stammt von Frobenius (vgl. das Zitat auf S. 81 u. 88).

Von Lehrbüchern verweisen wir auf: Muth, *Theorie und Anwendung der Elementarteiler*, Leipzig 1899, S. 20, Kronecker, 20. u. 21. Vorlesung üb. *Determinanten*, Weber, *Algebra* **2**, 163.

§ 7. Bilineare Formen und höhere komplexe Zahlen.

Die Lehre von den höheren komplexen Zahlen, die bereits im ersten Kapitel vom allgemeinen arithmetischen Standpunkt aus behandelt wurde, ist mit der Theorie der bilinearen Formen auf das engste verknüpft.

Die n^2 Matrices

$$F_{ik} \qquad (i, k = 1, 2, \ldots, n)$$

vom n^{ten} Grade, von denen jede nur in der i^{ten} Zeile an k^{ter} Stelle die 1, sonst lauter Nullen enthält oder die ihnen entsprechenden bilinearen Formen $x_i y_k$, sind offenbar linear unabhängig (vgl. S. 85). Für ihre Komposition gelten die Relationen:

$$(1)\quad F_{ik}F_{jl} = \delta_{kj}F_{il}\,(\delta_{kk} = 1, \delta_{kj} = 0, \text{ falls } k \gtrless j) \qquad (i, k, j, l = 1, 2, \ldots, n).$$

Mittels der n^2 Formen F_{ik} kann jede bilineare Form

$$A = \sum_{i=1}^{i=n} \sum_{k=1}^{k=n} a_{ik} x_i y_k$$

in der Form

$$A = \sum_{i=1}^{i=n} \sum_{k=1}^{k=n} a_{ik} F_{ik}$$

dargestellt werden.

Die n^2 Größen F_{ik} lassen sich als die n^2 unabhängigen Einheiten eines komplexen Zahlensystems (vgl. Kap. I, § 4) ansehen. Hierbei werden die „Einheitsprodukte“ durch die Relationen (1) bestimmt; diese sind so beschaffen, daß die Multiplikation assoziativ ist.

Jede bilineare Form

$$A = \sum_{i=1}^{i=n} \sum_{k=1}^{k=n} a_{ik} F_{ik}$$

kann demnach als komplexe aus n^2 unabhängigen Einheiten gebildete Zahl aufgefaßt werden.

Wird die bilineare Form

$$B = \sum_{i=1}^{i=n} \sum_{k=1}^{k=n} b_{ik} F_{ik}$$

ebenfalls als komplexe Zahl interpretiert und multipliziert man A

und B nach den Multiplikationsgesetzen für komplexe Zahl so erhält man mit Benutzung von (1):

$$AB = \sum_{i=1}^{i=n}\sum_{k=1}^{k=n}\sum_{j=1}^{j=n}\sum_{l=1}^{l=n} a_{ik} b_{jl} F_{ik} F_{jl}$$

$$= \sum_{i=1}^{i=n}\sum_{k=1}^{k=n}\sum_{j=1}^{j=n}\sum_{l=1}^{l=n} a_{ik} b_{jl} \delta_{kj} F_{il}$$

$$= \sum_{i=1}^{i=n}\sum_{l=1}^{l=n}\sum_{s=1}^{s=n} a_{is} b_{sl} F_{il} = \sum_{i=1}^{i=n}\sum_{l=1}^{l=n} c_{il} F_{il},$$

falls $$c_{il} = \sum_{s=1}^{s=n} a_{is} b_{sl} \qquad (i, l = 1, 2, \ldots, n)$$

gesetzt wird. Deutet man die komplexe Zahl

$$C = \sum_{i=1}^{i=n}\sum_{l=1}^{l=n} c_{il} F_{il}$$

als bilineare Form, so ist C nichts anderes als das Produkt AB der zwei bilinearen Formen A und B. *Die Komposition bilinearer Formen n^{ten} Grades erscheint demnach als identisch mit der Multiplikation besonderer komplexer Zahlen mit n^2 Einheiten.*

$G_1, G_2, \ldots, G_\varrho$ seien ϱ linear unabhängige bilineare Formen oder Matrices gleichen Grades n. Von ihnen wird vorausgesetzt, sie sollen die Eigenschaft haben, daß sich ihre Produkte

$$G_i G_k \qquad (i, k = 1, 2, \ldots, \varrho)$$

linear und homogen durch die ϱ Grundformen selbst darstellen lassen. Es sei also:

$$(2) \qquad G_i G_k = \lambda_{ik}^{(1)} G_1 + \lambda_{ik}^{(2)} G_2 + \cdots + \lambda_{ik}^{(\varrho)} G_\varrho \qquad (i, k = 1, 2, \ldots, \varrho),$$

wobei die Größen $\lambda_{ik}^{(r)}$ numerische Konstanten bedeuten. Infolge des assoziativen Gesetzes bei der Multiplikation bilinearer Formen besteht die Gleichung:

$$(G_i G_k) G_m = G_i (G_k G_m).$$

Aus ihr folgt, wenn man für die Produkte ihre Werte nach (2) setzt und die vorausgesetzte lineare Unabhängigkeit der Grundformen beachtet, daß:

$$(3) \qquad \sum_{s=1}^{s=\varrho} \lambda_{ik}^{(s)} \lambda_{sm}^{(r)} = \sum_{s=1}^{s=\varrho} \lambda_{is}^{(r)} \lambda_{km}^{(s)} \qquad (i, k, r, m = 1, 2, \ldots, \varrho)$$

wird. Aus diesen Gleichungen erhellt (vgl. Kap. I, § 4):

Sind $a_1, a_2, \ldots, a_\varrho$ irgendwelche numerische Konstante[1]*), so kann jede aus den Grundformen $G_1, G_2, \ldots, G_\varrho$ gebildete bilineare Form $a_1 G_1 + a_2 G_2 + \cdots + a_\varrho G_\varrho$ als eine aus den „Einheiten" $G_1, G_2, \ldots, G_\varrho$ gebildete komplexe Zahl angesehen werden.*

Umgekehrt wollen wir zeigen, wie jeder höheren komplexen Zahl eine bilineare Form zugeordnet werden kann. Sei eine aus den ϱ unabhängigen Einheiten $e_1, e_2, \ldots, e_\varrho$ gebildete komplexe Zahl $a_1 e_1 + a_2 e_2 + \cdots + a_\varrho e_\varrho$ gegeben. Für die „Einheitsprodukte" mögen die Formeln

$$e_i e_k = \lambda_{ik}^{(1)} e_1 + \lambda_{ik}^{(2)} e_2 + \cdots + \lambda_{ik}^{(\varrho)} e_\varrho \qquad (i, k = 1, 2, \ldots, \varrho)$$

gelten, wobei die Größen $\lambda_{ik}^{(r)}$ den Gleichungen (3) genügen.

Man konstruiere die ϱ bilinearen Formen oder Matrices ϱ^{ten} Grades:

$$G_\alpha = \sum_{r=1}^{r=\varrho} \sum_{m=1}^{m=\varrho} \lambda_{\alpha m}^{(r)} x_r y_m \qquad (\alpha = 1, 2, \ldots, \varrho).$$

Betrachtet man die Produkte $G_i G_k$ für $i, k = 1, 2, \ldots, \varrho$, so wird:

$$\begin{aligned} G_i G_k &= \sum_{r=1}^{r=\varrho} \sum_{m=1}^{m=\varrho} \sum_{s=1}^{s=\varrho} \lambda_{is}^{(r)} \lambda_{km}^{(s)} x_r y_m \\ &= \sum_{r=1}^{r=\varrho} \sum_{m=1}^{m=\varrho} \sum_{s=1}^{s=\varrho} \lambda_{ik}^{(s)} \lambda_{sm}^{(r)} x_r y_m \\ &= \sum_{s=1}^{s=\varrho} \lambda_{ik}^{(s)} G_s . \end{aligned}$$

Setzt man

$$e_i = G_i \qquad (i = 1, 2, \ldots, \varrho),$$

so läßt sich jede aus den ϱ Einheiten $e_1, e_2, \ldots, e_\varrho$ gebildete komplexe Zahl $a_1 e_1 + a_2 e_2 + \cdots + a_\varrho e_\varrho$ als eine aus den bi-

---

1) Die Größen $a_1, a_2, \ldots, a_\varrho$ brauchen nicht, wie im Kap. I, § 4 vorausgesetzt wurde, reell zu sein sie können auch komplexe Werte besitzen.

linearen Grundformen $G_1, G_2, \ldots, G_\varrho$ *gebildete bilineare Form oder Matrix* ϱ^{ten} *Grades* $a_1 G_1 + a_2 G_2 + \cdots + a_\varrho G_\varrho$ *auffassen.*

Die so gewonnenen Matrices $G_1, G_2, \ldots, G_\varrho$ *werden nicht notwendig linear unabhängig sein.* Dies lehrt beispielsweise das aus zwei Einheiten e_1 und e_2 gebildete Zahlensystem mit der Multiplikationsregel:

$$e_1 e_1 = e_1,\ e_2 e_1 = -e_1,\ e_1 e_2 = e_2,\ e_2 e_2 = -e_2.$$

Für dieses Zahlensystem wird G_1 die Einheitsform

$$E = x_1 y_1 + x_2 y_2, \quad G_2 = -E.$$

Damit die ϱ Matrices $G_1, G_2, \ldots, G_\varrho$ linear abhängig sind, ist notwendig und hinreichend, daß die ϱ^2 Gleichungen:

$$c_1 \lambda_{1m}^{(r)} + c_2 \lambda_{2m}^{(r)} + \cdots + c_\varrho \lambda_{\varrho m}^{(r)} = 0 \qquad (r, m = 1, 2, \ldots, \varrho)$$

nicht nur die Lösungen $c_1 = c_2 = \cdots = c_\varrho = 0$ besitzen. Bedeuten $a_1, a_2, \ldots, a_\varrho$ willkürliche numerische Konstanten, so erhält man hieraus die ϱ Gleichungen:

$$c_1 \sum_{\alpha=1}^{\alpha=\varrho} a_\alpha \lambda_{1\alpha}^{(r)} + c_2 \sum_{\alpha=1}^{\alpha=\varrho} a_\alpha \lambda_{2\alpha}^{(r)} + \cdots + c_\varrho \sum_{\alpha=1}^{\alpha=\varrho} a_\alpha \lambda_{\varrho\alpha}^{(r)} = 0 \qquad (r = 1, 2, \ldots, \varrho).$$

Aus diesen ϱ Gleichungen schließt man: Sind die ϱ bilinearen Formen $G_1, G_2, \ldots, G_\varrho$ in linearer Dependenz, so muß die Determinante:

$$|\Delta| = \left| \sum_{\alpha=1}^{\alpha=\varrho} a_\alpha \lambda_{i\alpha}^{(r)} \right| \qquad (i, r = 1, 2, \ldots, \varrho)$$

für jede Wahl der Größen $a_1, a_2, \ldots, a_\varrho$ verschwinden.

Verschwindet diese Determinante nicht identisch, so ist dies eine ausreichende Bedingung für die lineare Unabhängigkeit der bilinearen Formen $G_1, G_2, \ldots, G_\varrho$.

An Stelle der bilinearen Formen $G_1, G_2, \ldots, G_\varrho$ kann man auch die Formen:

$$H_\alpha = \sum_{r=1}^{r=\varrho} \sum_{m=1}^{m=\varrho} \lambda_{m\alpha}^{(r)} x_m y_r \qquad (\alpha = 1, 2, \ldots, \varrho)$$

einführen. Bildet man die Produkte $H_i H_k$ für $i, k = 1, 2, \ldots, \varrho$, so wird

$$H_i H_k = \sum_{s=1}^{s=\varrho} \lambda_{ik}^{(s)} H_s.$$

Setzt man $e_i = H_i$, *so läßt sich jede aus den* ϱ *Einheiten* $e_1, e_2, \ldots, e_\varrho$ *gebildete komplexe Zahl* $a_1 e_1 + a_2 e_2 + \cdots + a_\varrho e_\varrho$ *als eine bilineare Form* $a_1 H_1 + a_2 H_2 + \cdots + a_\varrho H_\varrho$ *interpretieren. Setzt man voraus, daß die Determinante:*

$$|\Gamma| = \left| \sum_{\alpha=1}^{\alpha=\varrho} a_\alpha \lambda_{\alpha i}^{(r)} \right| \qquad (i, r = 1, 2, \ldots, \varrho)$$

nicht für jede Wahl der Größen $a_1, a_2, \ldots, a_\varrho$ *verschwindet, so sind die bilinearen Formen* $H_1, H_2, \ldots, H_\varrho$ *linear unabhängig.*

Die Determinante $|\Gamma|$ ist offenbar die Determinante der bilinearen Form

$$a_1 G_1 + a_2 G_2 + \cdots + a_\varrho G_\varrho.$$

Das Nichtverschwinden der Determinante $|\Gamma|$ ist also damit gleichbedeutend, daß das Formensystem

$$a_1 G_1 + a_2 G_2 + \cdots + a_\varrho G_\varrho$$

bilineare Formen mit nichtverschwindenden Determinanten enthält. Das Nichtverschwinden der früher aufgetretenen Determinante $|\Delta|$ besagt, das das Formensystem

$$a_1 H_1 + a_2 H_2 + \cdots a_\varrho H_\varrho$$

auch Formen mit nichtverschwindenden Determinanten besitzt.

Mit Hilfe der bilinearen Formen G_α und H_α kann man die ϱ^4 gewöhnlichen Gleichungen (3) in ϱ^2 symbolische zusammenfassen, nämlich:

$$G_i H_m' = H_m' G_i \qquad (i, m = 1, 2, \ldots, \varrho),$$

wobei

$$H_m' = \sum_{r=1}^{r=\varrho} \sum_{k=1}^{k=\varrho} \lambda_{km}^{(r)} x_r y_k$$

die zu H_m transponierte Form ist. Hieraus folgt:

Sind $a_1, a_2, \ldots, a_\varrho$ *und* $b_1, b_2, \ldots, b_\varrho$ *beliebige Zahlen, so sind die zwei Matrices* $a_1 G_1 + a_2 G_2 + \cdots + a_\varrho G_\varrho$ *und* $b_1 H_1' + b_2 H_2' + \cdots + b_\varrho H_\varrho'$ *stets vertauschbar.*

Wir machen die für das Folgende wesentliche Voraussetzung, daß sowohl das Formensystem $a_1 H_1 + a_2 H_2 + \cdots + a_\varrho H_\varrho$ als auch das Formensystem $a_1 G_1 + a_2 G_2 + \cdots + a_\varrho G_\varrho$ bilineare Formen mit nichtverschwindenden Determinanten enthalten. Von

diesen zwei Bedingungen ist keine eine Folge der anderen, wie das oben angegebene Beispiel lehrt. Für dieses ist

$$H_1 = x_1 y_1 - x_2 y_1,\ H_2 = x_1 y_2 - x_2 y_2;$$
$$G_1 = x_1 y_1 + x_2 y_2,\ G_2 = -x_1 y_1 - x_2 y_2,$$

die Determinante $|a_1 H_1 + a_2 H_2|$ hat den Wert Null, während die Determinante $|a_1 G_1 + a_2 G_2|$ nicht identisch verschwindet. Die Voraussetzung, daß die zwei Determinanten:

$$|a_1 G_1 + a_2 G_2 + \cdots + a_\varrho G_\varrho| = |I|$$

und

$$|a_1 H_1 + a_2 H_2 + \cdots + a_\varrho H_\varrho| = |\Delta|$$

nicht identisch verschwinden, involviert, daß sowohl die Formen $H_1, H_2, \ldots, H_\varrho$ als auch $G_1, G_2, \ldots, G_\varrho$ linear unabhängig sein sollen.

Wir beschränken uns im folgenden auf die Betrachtung jener bilinearen Formen der Schar $a_1 G_1 + a_2 G_2 + \cdots + a_\varrho G_\varrho$, bei denen die numerischen Konstanten $a_1, a_2, \ldots, a_\varrho$ so gewählt sind, daß die Determinante $|a_1 G_1 + a_2 G_2 + \cdots + a_\varrho G_\varrho|$ nicht verschwindet. Sei $A = a_1 G_1 + a_2 G_2 + \cdots + a_\varrho G_\varrho$ eine solche Form. Wir suchen eine bilineare Form

$$U = u_1 G_1 + u_2 G_2 + \cdots + u_\varrho G_\varrho,$$

wobei $u_1, u_2, \ldots, u_\varrho$ numerische Konstanten sein sollen, zu bestimmen, daß das Produkt $AU = A$ wird. Die symbolische Gleichung $AU = A$ besagt:

$$(a_1 G_1 + a_2 G_2 + \cdots + a_\varrho G_\varrho) \cdot (u_1 G_1 + u_2 G_2 + \cdots u_\varrho G_\varrho) =$$
$$= \sum_{\alpha=1}^{\alpha=\varrho} \sum_{k=1}^{k=\varrho} a_\alpha u_k G_\alpha G_k = a_1 G_1 + a_2 G_2 + \cdots + a_\varrho G_\varrho.$$

Beachtet man die Gleichungen (2) und die auf Grund unserer Voraussetzungen stattfindende lineare Unabhängigkeit von $G_1, G_2, \ldots, G_\varrho$, so ergibt sich, daß

$$u_1 \sum_{\alpha=1}^{\alpha=\varrho} a_\alpha \lambda_{\alpha 1}^{(s)} + u_2 \sum_{\alpha=1}^{\alpha=\varrho} a_\alpha \lambda_{\alpha 2}^{(s)} + \cdots + u_\varrho \sum_{\alpha=1}^{\alpha=\varrho} a_\alpha \lambda_{\alpha\varrho}^{(s)} = a_s \quad (s = 1, 2, \ldots, \varrho)$$

wird. Da die Determinante:

$$|A| = |a_1 G_1 + a_2 G_2 + \cdots + a_\varrho G_\varrho| = \left| \sum_{\alpha=1}^{\alpha=\varrho} a_\alpha \lambda_{\alpha i}^{(r)} \right| \quad (i, r = 1, 2, \ldots, \varrho)$$

von Null verschieden ist, haben die zuletzt angegebenen ϱ Gleichungen endliche, eindeutig bestimmte Lösungen $u_1, u_2, \ldots, u_\varrho$. U ist offenbar die Einheitsform. Wir haben also bewiesen: *das Formensystem* $a_1 G_1 + a_2 G_2 + \cdots + a_\varrho G_\varrho$ *enthält stets die Einheitsform* E.

Sei $A = a_1 G_1 + a_2 G_2 + \cdots a_\varrho G_\varrho$ wiederum eine bilineare Form von nichtverschwindender Determinante, so kann man innerhalb des aus $G_1, G_2, \ldots, G_\varrho$ gebildeten Formensystemes eine Form $V = v_1 G_1 + v_2 G_2 + \cdots + v_\varrho G_\varrho$ zu bestimmen suchen, so daß $v_1, v_2, \ldots, v_\varrho$ numerische Konstanten bedeuten und $AV = E$ wird. Aus dieser symbolischen Gleichung gehen die ϱ gewöhnlichen Gleichungen:

$$v_1 \sum_{\alpha=1}^{\alpha=\varrho} a_\alpha \lambda_{\alpha 1}^{(s)} + v_2 \sum_{\alpha=1}^{\alpha=\varrho} a_\alpha \lambda_{\alpha 2}^{(s)} + \cdots + v_\varrho \sum_{\alpha=1}^{\alpha=\varrho} a_\alpha \lambda_{\alpha\varrho}^{(s)} = u_s \quad (s = 1, 2, \ldots, \varrho)$$

hervor. Hierbei sind $u_1, u_2, \ldots, u_\varrho$ die durch

$$E = u_1 G_1 + u_2 G_2 + \cdots u_\varrho G_\varrho$$

festgelegten numerischen Konstanten. Da die Determinante

$$|A| = \left| \sum_{\alpha=1}^{\alpha=\varrho} a_\alpha \lambda_{\alpha i}^{(r)} \right| \qquad (i, r = 1, 2, \ldots, \varrho)$$

von Null verschieden ist, sind die ϱ Gleichungen lösbar und bestimmen eindeutige, endliche Werte $v_1, v_2, \ldots, v_\varrho$. V ist offenbar die reziproke Matrix von A.

Wir haben das Resultat:

Verschwinden die zwei Determinanten

$$|a_1 G_1 + a_2 G_2 + \cdots + a_\varrho G_\varrho| \text{ und } |a_1 H_1 + a_2 H_2 + \cdots + a_\varrho H_\varrho|$$

nicht für alle Werte $a_1, a_2, \ldots, a_\varrho$, so ist die Gesamtheit $\mathfrak{G}$ aller Matrices $a_1 G_1 + a_2 G_2 + \cdots + a_\varrho G_\varrho$ von nichtverschwindenden Determinanten bei der Produktbildung nicht nur in sich abgeschlossen, sondern $\mathfrak{G}$ enthält auch das Einheitselement und zu jeder Matrix die reziproke. *Die Gesamtheit aller in* $\mathfrak{G}$ *enthaltenen Matrices bildet also (Kap. III, § 1) eine Gruppe.*

Verschwinden die zwei erwähnten Determinanten nicht identisch, so bildet auch die Gesamtheit $\mathfrak{H}$ *aller Matrices*

$$a_1 H_1 + a_2 H_2 + \cdots + a_\varrho H_\varrho$$

von nichtverschwindenden Determinanten eine Gruppe $\mathfrak{H}$.

Jede Matrix der Gruppe $\mathfrak{H}$ ist mit jeder Matrix der Gruppe $\mathfrak{G}'$, die von den transponierten Matrices der Gruppe $\mathfrak{G}$ gebildet wird, vertauschbar. ($\mathfrak{H}$ und $\mathfrak{G}'$ sind in Lies Terminologie reziproke Gruppen. Lie-Engel, *Theorie der Transformationsgruppen*, Leipzig 1893, **3**, 751 u. 779.)

In der Sprache der höheren komplexen Zahlen besagt das Nichtverschwinden der zwei Determinanten, daß im Gebiet der behandelten höheren komplexen Zahlen die Umkehrung der Multiplikation als vordere und als hintere Division im allgemeinen eindeutig möglich ist.

Verschwinden eine oder beide der Determinanten $|\Gamma|$ und $|\Delta|$, so führe man noch zu den ϱ Einheiten $e_1, e_2, \ldots, e_\varrho$ eine weitere Einheit e_0 ein. Für die Multiplikation mit dieser Einheit mögen die Regeln

$$e_0 e_i = e_i e_0 = e_i \qquad (i = 0, 1, 2, \ldots, \varrho)$$

gelten. Betrachtet man die $\varrho + 1$ Einheiten $e_0, e_1, e_2, \ldots, e_\varrho$, so sind die Konstanten $\lambda_{ik}^{(i)}$ für sie auf folgende Weise bestimmt:

$$\lambda_{i0}^{(i)} = \lambda_{0i}^{(i)} = 1 \qquad (i = 0, 1, 2, \ldots, \varrho),$$

alle weiteren $\lambda_{ik}^{(i)} = 0$, bei denen ein Index 0 ist, und, falls keine der Zahlen i, k, s gleich Null ist, sind $\lambda_{ik}^{(i)}$ die durch die Gleichungen (2) bestimmten Größen.

Definiert man:

$$G_\alpha = \sum_{r=0}^{r=\varrho} \sum_{m=0}^{m=\varrho} \lambda_{\alpha m}^{(r)} x_r y_m, \qquad H_\alpha = \sum_{r=0}^{r=\varrho} \sum_{m=0}^{m=\varrho} \lambda_{m\alpha}^{(r)} x_m y_r \qquad (\alpha = 0, 1, 2, \ldots, \varrho),$$

so ist

$$G_0 = H_0 = x_0 y_0 + x_1 y_1 + \cdots + x_\varrho y_\varrho.$$

Mithin verschwindet keine der zwei Determinanten

$$|a_0 G_0 + a_1 G_1 + \cdots + a_\varrho G_\varrho| \text{ und } |a_0 H_0 + a_1 H_1 + \cdots + a_\varrho H_\varrho|$$

für alle Werte $a_0, a_1, a_2, \ldots, a_\varrho$; denn für

$$a_0 = 1, \; a_1 = a_2 = \cdots = a_\varrho = 0$$

nehmen diese Determinanten den Wert 1 an.

Wenden wir die obigen Sätze an, so erhalten wir das Resultat:

Ist eine der zwei Determinanten

$$|\Delta| = \left| \sum_{\alpha=1}^{\alpha=\varrho} a_\alpha \lambda_{i\alpha}^{(r)} \right| \qquad (i, r = 1, 2, \ldots, \varrho)$$

und

$$|\Gamma| = \left| \sum_{\alpha=1}^{\alpha=\varrho} a_\alpha \lambda_{\alpha i}^{(r)} \right| \qquad (i, r = 1, 2, \ldots, \varrho)$$

oder beide Null, so bildet die Gesamtheit $\mathfrak{G}_0$ *aller Matrices*

$$a_0 G_0 + a_1 G_1 + \cdots + a_\varrho G_\varrho$$

vom $\varrho + 1^{ten}$ *Grade mit nichtverschwindenden Determinanten eine Gruppe. Eben dieses trifft für die Gesamtheit* $\mathfrak{H}_0$ *aller Matrices*

$$a_0 H_0 + a_1 H_1 + \cdots + a_\varrho H_\varrho$$

vom $\varrho + 1^{ten}$ *Grade mit nichtverschwindenden Determinanten zu.*

Auf die Deutung höherer komplexer Zahlen durch Matrices hat bereits (1858) Cayley (*Coll. math. papers* **2**, 491) in seiner grundlegenden Arbeit über Matrices (vgl. den voraufgegangenen § 6) für den Fall der Quaternionen hingewiesen. Die vier Matrices zweiten Grades:

$$i_0 = \left\| \begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix} \right\|, \quad i_1 = \left\| \begin{matrix} 0 & i \\ i & 0 \end{matrix} \right\|, \quad i_2 = \left\| \begin{matrix} i & 0 \\ 0 & -i \end{matrix} \right\|, \quad i_3 = \left\| \begin{matrix} 0 & 1 \\ -1 & 0 \end{matrix} \right\|$$

verhalten sich bei ihrer Komposition wie die Quaternioneneinheiten (vgl. S. 16). Die weitere Deutung komplexer Zahlen durch Matrices verdankt man C. S. Peirce (*Am. J. math.* **4**, 221 (1881)) und Ed. Weyr (*Prager Ber.* (1887)). Die Verknüpfung höherer komplexer Zahlen mit Gruppen geht auf Poincaré (*C. R.* **99** (1884), 740) und Study (*Monatsh. f. Math.* **1**, 283 (1890)) zurück. Wir nennen ferner: Scheffers, *Math. Ann.* **39**, 293 (1891), ebenda, **41**, 601 (1893), Molien, ebenda, **41**, 83 (1893), **42**, 308 (1893). Wegen weiterer Litteratur über den Zusammenhang höherer komplexer Zahlen und bilinearer Formen sei auf Studys Artikel in der *Enzyklopädie der math. Wiss.* **1**, 147 *„Theorie der gemeinen und höheren komplexen Größen"* verwiesen. Zur Ergänzung führen wir noch die rein algebraischen Arbeiten von Frobenius, *Theorie der hyperkomplexen Größen* (*Sitzungsb. d. Berl. Akad.* (1903), 504 u. 634) an. Von Lehrbüchern vgl. man: Lie-Scheffers, *Vorl. über kontinuierliche Gruppen*, Leipzig 1893, S. 610.

§ 8. Die Theorie der Elementarteiler und die bilinearen Formen.

Sind A und B zwei quadratische Matrices n^{ten} Grades mit den Elementen a_{ik} und b_{ik}, so ist unter $A + \varrho B$ (vgl. S. 81) die Matrix mit den Elementen $a_{ik} + \varrho b_{ik}$ zu verstehen. Die Behandlung einer solchen Matrix $A + \varrho B$, deren Elemente ganze Funktionen ersten Grades eines Parameters ϱ sind, führte Weierstraß (*Zur Theorie der quadratischen und bilinearen Formen, Monatsb. d. Berl. Akad.* (1868), mit Veränderungen abgedruckt in *Ges. Werken* 2, 19) auf seine berühmte Theorie der *Elementarteiler.*

Unser Ausgangspunkt soll im folgenden ein rein *zahlentheoretischer* sein; hierdurch erhält man die Begriffe der Theorie wohl am klarsten und durchsichtigsten. Alsdann legen wir die Allgemeinheit dar, die diesen Ideenbildungen zukommt, und gelangen erst am Schlusse durch Spezialisierung zu dem Standpunkt, den Weierstraß eingenommen hat.

Wir betrachten die quadratische Matrix:

$$D = \| d_{ik} \| \qquad (i, k = 1, 2, \ldots, n)$$

vom Grad n. Ihre Elemente d_{ik} sollen ausnahmslos ganze Zahlen, sowohl positiv als negativ, einschließlich der Null sein. Die Matrix D habe den Rang l, d. h. alle ihre Unterdeterminanten vom $l + 1^{\text{ten}}$, jedoch nicht sämtliche vom l^{ten} Grad seien Null.

σ bedeute im folgenden eine ganze positive Zahl, die $\leqq l$ ist, mit T_σ bezeichnen wir (den positiv gewählten) größten gemeinsamen Teiler sämtlicher Unterdeterminanten σ^{ten} Grades D_σ von D. Ist $l = n$, so ist $T_l = T_n$ und, vom Vorzeichen abgesehen, die Determinante der Matrix D. Ist $l < n$, so setze man $T_{l+1} = T_{l+2} = \cdots = T_n = 0$.

Definition I: Die Größen $T_1, T_2, \ldots, T_n$ heißen der erste, zweite, . . ., n^{te} Determinantenteiler des Systems der d_{ik} oder der Determinante D.

Jede σreihige Unterdeterminante D_σ ist eine lineare homogene Funktion von Unterdeterminanten $D_{\sigma-1}$ des $\sigma - 1^{\text{ten}}$ Grades. Mithin ist der größte gemeinsame Teiler T_σ aller D_σ durch den größten gemeinsamen Teiler $T_{\sigma-1}$ aller $D_{\sigma-1}$ teilbar.

Lehrsatz I: Der Quotient

$$\frac{T_\sigma}{T_{\sigma-1}} \qquad (\sigma = 2, 3, \ldots, l)$$

des σ^{ten} Determinantenteilers durch den $\sigma - 1^{ten}$ ist eine ganze Zahl.

Definition II: Die Größen

$$E_1 = T_1,\ E_\sigma = \frac{T_\sigma}{T_{\sigma-1}} \quad (\sigma = 2, 3, \ldots, l), \qquad E_{l+\mu} = 0 \quad (\mu = 1, 2, \ldots, n-l)$$

heißen die zusammengesetzten Elementarteiler des Systems der d_{ik} oder der Determinante D. E_1 heißt der erste, E_2 der zweite, ..., E_l der l^{te}, E_{l+1} der $l+1^{te}$, ..., E_n der n^{te} Elementarteiler. Die Anzahl der nichtverschwindenden Elementarteiler gibt ebenso wie die der nichtverschwindenden Determinantenteiler den Rang von D an.

Die Bezeichnung der Größen $E_1, E_2, \ldots, E_n$ als zusammengesetzte Elementarteiler ist Muths Werk *„Theorie und Anwendung der Elementarteiler"*, Leipzig 1899, S. 13, entlehnt. Der Name „σ^{ter} Elementarteiler" ist von Frobenius (*Journ. f. Math.* **86**, 148 (1879)) eingeführt worden; die Bezeichnung zusammengesetzter Elementarteiler wird jedoch in seiner grundlegenden Arbeit (vgl. a. a. O. S. 163) in anderer Bedeutung als hier verwandt. Die Größen $E_1, E_2, \ldots, E_n$ finden sich bereits (1861), ehe Weierstraß seine klassische Arbeit, aus welcher der Name Elementarteiler stammt, publiziert hatte, in Henry J. St. Smiths zahlentheoretischen Untersuchungen über lineare Gleichungen mit ganzzahligen Koeffizienten (Smith, *Coll. math. papers* **I**, 391).

Von Smith stammen auch schon folgende Sätze:

Lehrsatz II: In der Reihe der zusammengesetzten Elementarteiler $E_1, E_2, \ldots, E_n$ ist jeder Elementarteiler E_σ durch den ihm voraufgehenden $E_{\sigma-1}$ $(\sigma = l, l-1, \ldots, 2)$ ohne Rest teilbar (Smith, a. a. O., S. 396).

Lehrsatz III: Die zusammengesetzten Elementarteiler lassen sich nicht nur als Quotienten

$$E_\sigma = \frac{T_\sigma}{T_{\sigma-1}}$$

größter gemeinsamer Teiler definieren, sondern sie selbst sind auch größte gemeinsame Teiler: Dividiert man jede Unterdeterminante σ^{ten} Grades D_σ von D durch den größten gemeinsamen Teiler der in D_σ enthaltenen Unterdeterminanten $\sigma - 1^{ten}$ Grades, so ist der größte gemeinsame Teiler aller dieser Quotienten gleich dem σ^{ten} Elementarteiler E_σ (Smith, a. a. O., S. 396).

Aus dem Lehrsatz II folgt: $E_l \geqq E_{l-1} \geqq E_{l-2} \cdots \geqq E_1$. Mithin ergibt sich: Sind die Werte der l zusammengesetzten, nichtverschwindenden Elementarteiler $E_l, E_{l-1}, \ldots, E_1$ in irgend welcher Reihenfolge gegeben, so braucht man sie nur ihrer Größe nach zu ordnen, um zu entscheiden, welcher der σ^{te} ist.

Lehrsatz IV: Komponiert man zwei oder mehrere ganzzahlige Matrices, so ist der σ^{te} Elementarteiler des Produktes durch den σ^{ten} Elementarteiler eines jeden Faktors des Produktes teilbar.

Aus dem Satz II folgt: Jeder zusammengesetzte Elementarteiler $E_{\sigma-1}$ $(\sigma = l, l-1, \ldots, 2)$ ist nur durch solche Primzahlen teilbar, die auch E_σ ohne Rest teilen. Hieraus ergibt sich:

Lehrsatz V: Ist p^{e_σ} die höchste Potenz einer Primzahl p, die in E_σ $(\sigma = l, l-1, \ldots, 1)$ ohne Rest aufgeht, so genügen die Zahlen e_σ den Ungleichungen

$$e_l \geqq e_{l-1} \geqq e_{l-2} \cdots \geqq e_1.$$

Diese Fundamentaleigenschaft ist von Cayley (1855) (*Journ. f. Math.* **50**, 314, *Coll. math. papers* **2**, 217) angegeben worden.

Definition III: Ist p^{e_σ} die höchste in E_σ $(\sigma = l, l-1, \ldots, 1)$ enthaltene Potenz einer Primzahl p, so heißt, falls e_σ von Null verschieden ist, jeder Faktor p^{e_σ} nach Weierstraß ein Elementarteiler oder eine elementare Invariante des Systemes d_{ik} oder der Determinante D. Die Primzahl p wird die Grundzahl oder Basis, e_σ der Exponent oder Grad des Elementarteilers p^{e_σ} genannt.

Es empfiehlt sich oft, die Elementarteiler p^{e_σ} mit Frobenius (*Journ. f. Math.* **86**, 162) zum Unterschied von den zusammengesetzten Elementarteilern als *einfache* Elementarteiler zu bezeichnen. Spricht man von Elementarteilern ohne Zusatz, so meint man die eben definierten einfachen Weierstraßschen (Kronecker verwendet in den *Monatsb. d. Berl. Akad.* (1874), *Ges. Werke* **1**, 405 die Bezeichnung „einfacher Elementarteiler" in ganz anderem Sinn; nach ihm hätte man Elementarteiler mit dem Exponenten 1 so zu nennen).

Die drei durch die Definitionen I, II, III eingeführten Größensysteme, nämlich die Determinantenteiler T_σ, die zusammengesetzten Elementarteiler E_σ und die einfachen Elementarteiler bestimmen sich gegenseitig eindeutig.

Die zusammengesetzten Elementarteiler E_σ sind durch die Definition II aus den Determinantenteilern abgeleitet worden. Sämtliche einfache oder Weierstraßsche Elementarteiler erhält man durch die Zerlegung des ersten, zweiten usw. bis l^{ten} Elementarteilers in die (von Null verschiedenen) Potenzen verschiedener Primfaktoren. Ist $E_\sigma = p^{e_\sigma} \cdot q^{e'_\sigma} \cdot r^{e''_\sigma} \cdots$, wobei $p, q, r, \ldots$ lauter verschiedene Primzahlen bedeuten, so sind die Größen

$$p^{e_\sigma}, \; q^{e'_\sigma}, \; r^{e''_\sigma}, \; \ldots \qquad (\sigma = 1, 2, \ldots, l)$$

die sämtlichen einfachen oder Weierstraßschen Elementarteiler.

Aus den zusammengesetzten Elementarteilern ergeben sich die Determinantenteiler T auf Grund der Relationen:

$$T_\sigma = E_\sigma \cdot E_{\sigma-1} \cdot E_{\sigma-2} \cdots E_1 \qquad (\sigma = 1, 2, \ldots, l),$$

$$T_{l+\mu} = E_{l+\mu} \qquad (\mu = 1, 2, \ldots, n-l).$$

Sind

$$p^{e_l}, \; p^{e_l - 1}, \; p^{e_l - 2}, \; \ldots, \; q^{e'_l}, \; q^{e'_l - 1}, \; q^{e'_l - 2}, \; \ldots$$

sämtliche einfache Elementarteiler von D — ihr Produkt muß T_l sein —, die nach fallenden Potenzen von $p, q, \ldots$ geordnet sind:

$$(e_l \geqq e_{l-1} \geqq e_{l-2} \geqq \cdots; \; e'_l \geqq e'_{l-1} \geqq e'_{l-2} \geqq \cdots, \cdots), \; .$$

so ist $E_n = E_{n-1} = \cdots = E_{l+1}$ durch den Rang von D bestimmt. E_l ist das Produkt der höchsten Potenzen sämtlicher einfacher Elementarteiler von D, also gleich $p^{e_l} \cdot q^{e'_l} \cdot \cdots$, E_{l-1} das Produkt der zweithöchsten Potenzen sämtlicher einfacher Elementarteiler, also gleich $p^{e_l - 1} \cdot q^{e'_l - 1} \cdot \cdots$, usw.

Alle diese Sätze sind *weitgehender Verallgemeinerung* fähig. Sie beruhen auf der eindeutigen Zerlegbarkeit der ganzen positiven Zahlen in Primfaktoren und der Existenz des größten gemeinsamen Teilers von zwei oder mehreren ganzen Zahlen. Ω sei ein *Rationalitätsbereich* oder *Zahlkörper* (vgl. Algebra). $\varrho_1, \varrho_2, \ldots, \varrho_k$ seien k Variable. Sind die Größen α Zahlen aus Ω, so heißen Summen von Gliedern der Form:

$$\alpha \varrho_1^{r_1} \varrho_2^{r_2} \ldots \varrho_k^{r_k},$$

wobei $r_1, r_2, \ldots, r_k$ ganze positive Zahlen, einschließlich der Null, bedeuten, ganze Funktionen in Ω. Diese lassen sich in *reduzible* (zerlegbare) und *irreduzible* (unzerlegbare) unterscheiden. Die ersteren sind als Produkte aus mehreren ganzen Funktionen in Ω darstellbar, die anderen nicht. Die reduziblen Funktionen lassen sich in eine endliche Anzahl irreduzibler Faktoren zerlegen, die selbst ganze Funktionen in Ω sind; jede reduzible Funktion bestimmt ihre irreduziblen Faktoren bis auf multiplikative Konstanten eindeutig. Hierauf gründet sich der Begriff des *größten gemeinschaftlichen Teilers* von zwei oder mehreren ganzen Funktionen in Ω; er ist bis auf einen konstanten Faktor eindeutig bestimmt (vgl. beispielsweise H. Weber, *Algebra* **2**, 563).

Die n^2 Elemente der Matrix

$$D = \| d_{ik} \| \qquad (i, k = 1, 2, \ldots, n)$$

seien ganze Funktionen in $\mathfrak{Q}$ von k Variablen. Auch hierfür läßt sich nach den obigen Angaben der größte gemeinsame Teiler T_σ sämtlicher Unterdeterminanten σ^{ten} Grades von D definieren. Die Determinantenteiler $T_1, T_2, \ldots, T_n$ sind, abgesehen von konstanten Faktoren, die hier die Rolle der Einheit spielen und für das Folgende als unwesentlich anzusehen sind, eindeutig bestimmt. Ebendasselbe gilt für die durch die Definition II eingeführten zusammengesetzten Elementarteiler. Anstatt der Primzahlen p treten im Falle der ganzen Funktionen in $\mathfrak{Q}$ die irreduziblen ganzen Funktionen von k Variablen mit Koeffizienten aus $\mathfrak{Q}$. Sämtliche einfache oder Weierstraßsche Elementarteiler erhält man durch die Zerlegung jedes zusammengesetzten Elementarteilers E_σ in die (von Null verschiedenen) Potenzen verschiedener irreduzibler Faktoren. Ist

$$E_\sigma = p^{e_\sigma} \cdot q^{e'_\sigma} \cdot r^{e''_\sigma} \cdot \ldots,$$

wobei die Basen $p, q, r, \ldots$ lauter voneinander verschiedene irreduzible ganze Funktionen mit Koeffizienten aus $\mathfrak{Q}$ sind, d. h. solche, die sich nicht nur um multiplikative Konstanten unterscheiden, so sind

$$p^{e_\sigma}, \; q^{e'_\sigma}, \; r^{e''_\sigma}, \ldots \qquad (\sigma = 1, 2, \ldots, l)$$

die sämtlichen einfachen oder Weierstraßschen Elementarteiler. *Alle früher angegebenen Lehrsätze behalten ihre unveränderte Gültigkeit, wenn man Matrices betrachtet, deren Elemente ganze Funktionen von k Variablen mit Koeffizienten aus einem Rationalitätsbereich $\mathfrak{Q}$ sind.*

D und F seien zwei Matrices mit je n^2 gleichartigen Elementen d_{ik} und f_{ik}, d. h. entweder seien beide Größensysteme d_{ik} und f_{ik} ganzzahlig oder beide ganze Funktionen von k Variablen mit Koeffizienten aus dem gleichen Rationalitätsbereich $\mathfrak{Q}$. Besteht zwischen F und D eine symbolische Gleichung $F = RDQ$, wobei R und Q Matrices mit *gleichartigen Elementen* wie denjenigen von F und D sein sollen, so soll auch bei dieser Beschränkung der Elemente von R und Q für F der Ausdruck *Vielfaches,* für D *Teiler* verwandt werden (zu vgl. S. 89). Waren die Elemente von R und Q willkürlich, so galt nur der Satz, daß der Rang des Vielfachen F stets gleich oder kleiner als der

des Teilers D ist. Bei der Beschränkung der Elemente von R und Q ergibt sich nach Satz IV:

Lehrsatz VI: Eine Matrix F ist nur dann ein Vielfaches einer Matrix D, wenn ihre zusammengesetzten Elementarteiler E_i' Vielfache der entsprechenden Elementarteiler E_i von D sind, also $E_i' = \gamma_i E_i$, wobei die γ_i ganze Größen (also ganze Zahlen oder ganze Funktionen) *einschließlich der Null sind.*

Dem Obigen entsprechend mögen zwei Matrices D und F mit gleichartigen Elementen *äquivalent* heißen, wenn sowohl F ein Vielfaches von D als auch D ein Vielfaches von F in dem engeren Sinne ist. Aus Satz VI folgt

Lehrsatz VII: Zwei Matrices D und F mit gleichartigen Elementen können nur dann äquivalent sein, wenn die entsprechenden zusammengesetzten Elementarteiler E_i und E_i' gleich sind.

Die in dem letzten Satz ausgesprochene Bedingung für die Aquivalenz zweier gleichartiger Matrices ist zwar *notwendig*, aber nicht *hinreichend*. Gegenüber zuweit gehenden Behauptungen in dieser Hinsicht scheint es angebracht, an einem Beispiel zu zeigen, daß *trotz der Gleichheit der Elementarteiler eine Matrix nicht Vielfaches einer anderen zu sein braucht.*

Sei D die Matrix

$$\begin{Vmatrix} \varrho_1\varrho_2 & 0 \\ 0 & 1 \end{Vmatrix}$$

und F die Matrix

$$\begin{Vmatrix} \varrho_1 & 0 \\ 0 & \varrho_2 \end{Vmatrix};$$

für F und D sind, falls ϱ_1 und ϱ_2 unabhängige Variable bedeuten und Ω der Bereich aller Zahlen ist, die zusammengesetzten Elementarteiler $E_1 = 1$, $E_2 = \varrho_1\varrho_2$ gleich. Angenommen F sei Vielfaches von D, nämlich $F = RDQ$, wobei die Elemente r_{ik} und q_{ik} der zwei Matrices R und Q ganze Funktionen der zwei Variablen ϱ_1, ϱ_2 sind. Aus $|D| = |F|$ folgt $|R| \cdot |Q| = 1$. Da das Produkt zweier ganzer Funktionen nur dann gleich 1 ist, falls sie Konstante sind, muß $|R| = \text{Konst.}$, $|Q| = \frac{1}{\text{Konst.}}$ sein. Mithin existiert $P = Q^{-1}$, und die Elemente p_{ik} von P sind wegen $|Q|^{-1} = \text{Konst.}$ ganze Funktionen von ϱ_1 und ϱ_2. Die symbolische Gleichung $FP = RD$ ergibt die vier gewöhnlichen Gleichungen:

$$\varrho_1 p_{11} = r_{11}\varrho_1\varrho_2, \quad \varrho_1 p_{12} = r_{12}, \quad \varrho_2 p_{21} = r_{21}\varrho_1\varrho_2, \quad \varrho_2 p_{22} = r_{22}.$$

Folglich wird die Determinante:

$$|R| = \begin{vmatrix} r_{11} r_{12} \\ r_{21} r_{22} \end{vmatrix} = \begin{vmatrix} r_{11} & \varrho_1 p_{12} \\ r_{21} & \varrho_2 p_{22} \end{vmatrix}.$$

Diese Determinante verschwindet für $\varrho_1 = \varrho_2 = 0$. Sie soll aber unabhängig von ϱ_1, ϱ_2 einen von Null verschiedenen konstanten Wert besitzen. Mithin ist die Annahme, daß F Vielfaches von D ist, widerlegt.[1])

Beschränkt man aber die Elemente d_{ik} und f_{ik} der zwei Matrices D und F darauf, *ganze Zahlen oder ganze Funktionen einer einzigen Variablen* ϱ *mit Koeffizienten aus* $\mathfrak{Q}$ zu sein, so sind die in den Sätzen VI und VII ausgesprochenen Bedingungen auch *ausreichend.* Im folgenden wird *stets ausnahmslos* vorausgesetzt, daß wir es mit Matrices, deren Elemente *ganzzahlig oder ganze Funktionen einer einzigen Variablen* ϱ *mit Koeffizienten aus* $\mathfrak{Q}$ sind, zu tun haben. Wir formulieren

Lehrsatz VI′: Eine Matrix F, deren Elemente ganze Zahlen oder ganze Funktionen einer einzigen Variablen ϱ *mit Koeffizienten aus* $\mathfrak{Q}$ *sind, ist dann und nur dann Vielfaches einer gleichartigen Matrix D, wenn ihre zusammengesetzten Elementarteiler* E_i' *Vielfache der entsprechenden Elementarteiler* E_i *von D sind, also* $E_i' = \gamma_i E_i$, *wobei* γ_i *ganze Größen* (also ganze Zahlen oder ganze Funktionen einer einzigen Variablen) *einschließlich Null sind.*

---

1) Das obige Beispiel läßt sich auch auf folgende Weise behandeln: Würde D Vielfaches von F sein, so müßte dies auch noch statthaben, wenn man die zwei unabhängigen Variablen ϱ_1 und ϱ_2 gleichsetzt. Sind F_0 und D_0 die sich aus F und D für $\varrho_1 = \varrho_2$ ergebenden Matrices, so hat

$$F_0 = \left\| \begin{matrix} \varrho_1 & 0 \\ 0 & \varrho_1 \end{matrix} \right\|$$

die Elementarteiler $E_1 = \varrho_1$, $E_2 = \varrho_1^2$, hingegen besitzt

$$D_0 = \left\| \begin{matrix} \varrho_1^2 & 0 \\ 0 & 1 \end{matrix} \right\|$$

die Elementarteiler $E_1 = 1$, $E_2 = \varrho_1^2$. Da der Elementarteiler $E_1 = 1$ von D_0 nicht Vielfaches des Elementarteilers $E_1 = \varrho_1$ von F_0 ist, kann nach Satz VI D_0 nicht Vielfaches von F_0, also D nicht Vielfaches von F sein. Daß bei Funktionen von zwei oder mehr Variablen die in den Sätzen VI und VII ausgesprochenen Bedingungen nicht ausreichen, hat seinen Grund darin, daß bei Funktionen von zwei und mehr Variablen der größte gemeinsame Teiler nicht alles den Funktionen „Gemeinsame" erschöpft.

Lehrsatz VII′: Die notwendige und zugleich hinreichende Bedingung für die Äquivalenz zweier gleichartiger Matrices mit Elementen, die ganze Zahlen oder ganze Funktionen einer einzigen Variablen mit Koeffizienten aus Ω sind, läßt sich in folgende drei gleichwertige Formen einkleiden: Übereinstimmung in den Determinantenteilern oder in den zusammengesetzten Elementarteilern oder im Rang und den einfachen Elementarteilern.

Hat die Determinante einer Matrix oder linearen homogenen Substitution mit ganzzahligen Koeffizienten den Wert ± 1, so heißt die Matrix oder Substitution *unimodular*. Hat man Matrices oder Substitutionen, deren Koeffizienten ganze Funktionen einer Variablen ϱ sind, so bezeichnet man diejenigen von ihnen, deren Determinanten von ϱ unabhängig und von Null verschieden sind, als *unimodular*. Man beweist, *daß für zwei nach Satz VII′ äquivalente Matrices D und F mit gleichartigen Koeffizienten in der symbolischen Gleichung $F = RDQ$, wobei R und Q Koeffizienten derselben Art wie F und D haben, die Matrices R und Q stets unimodular wählbar sind; für unimodulare Systeme R und Q existieren R^{-1} und Q^{-1} und sind ebenfalls unimodular. Es wird $D = R^{-1}FQ^{-1}$.*

Überträgt man (zu vgl. § 6) im Falle von Matrices mit ganzzahligen Elementen diese Sätze in die Sprache der Theorie der bilinearen Formen, so erhält man die Theoreme:

Lehrsatz VIII: Zwei bilineare Formen mit ganzzahligen Koeffizienten sind dann und nur dann äquivalent in dem Sinn, daß sie gegenseitig durch ganzzahlige (im allgemeinen nicht kogrediente) Substitutionen ineinander transformierbar sind, wenn die entsprechenden zusammengesetzten Elementarteiler der Matrices der Koeffizienten der beiden Formen gleich sind. Für die zusammengesetzten Elementarteiler kann man auch die Determinantenteiler oder den Rang und die einfachen Elementarteiler treten lassen. Im Falle der Äquivalenz können die zwei bilinearen Formen stets auch ganzzahlig unimodular ineinander übergeführt werden. Die Äquivalenz läßt sich auf rationalem Wege (durch Aufsuchung größter gemeinsamer Teiler) entscheiden, und die überführenden unimodularen Substitutionen lassen sich rational ermitteln.

Hieraus folgt *Lehrsatz IX: Eine gegebene bilineare Form mit ganzzahligen Koeffizienten, bei der die aus den Koeffizienten gebildete Matrix die zusammengesetzten, nichtverschwindenden Elementarteiler*

$$E_1, E_2, \ldots, E_t$$

hat, läßt sich durch ganzzahlige, unimodulare (im allgemeinen nicht kogrediente) Substitutionen für die zwei Variablenreihen in

$$E_1 x_1' y_1' + E_2 x_2' y_2' + \cdots + E_i x_i' y_i'$$

transformieren.

Die Sätze VIII und IX bleiben *unverändert gültig,* wenn man überall anstatt der ganzzahligen Koeffizienten *ganze Funktionen in einem beliebigen Rationalitätsbereich* Ω *einer einzigen Variablen* ϱ voraussetzt.

Also:

$$D = \sum_{i=1}^{i=n} \sum_{k=1}^{k=n} d_{ik} x_i y_k$$

und

$$F = \sum_{i=1}^{i=n} \sum_{k=1}^{k=n} f_{ik} x_i y_k$$

seien zwei bilineare Formen, deren Koeffizienten ganze Funktionen beliebigen Grades einer einzigen Variablen ϱ sind. Die Faktoren der einzelnen Potenzen von ϱ in den ganzen Funktionen d_{ik} und f_{ik} sollen einem Rationalitätsbereich Ω angehören. Die zwei bilinearen Formen heißen *äquivalent,* wenn sie durch Substitutionen ineinander überführbar sind, deren Koeffizienten ganze Funktionen von ϱ sind Zwei bilineare Formen der eben beschriebenen Art sind dann und nur dann äquivalent, wenn sie die gleichen zusammengesetzten Elementarteiler haben. Zwei solche äquivalente bilineare Formen lassen sich auch stets durch Substitutionen ineinander überführen, deren Koeffizienten ganze Funktionen in Ω der einen Variablen ϱ sind und deren Determinanten einen von ϱ unabhängigen konstanten Wert haben.

Setzt man die Elemente von D und F nicht als beliebige, sondern als *lineare Funktionen des Parameters* ϱ voraus, so kann man die Resultate noch verschärfen und ist bei dem Weierstraßschen Standpunkt, von dem wir zu Beginn des Paragraphen sprachen, angelangt. Die Elemente d_{ik} von D seien von der Form

$$d_{ik} = a_{ik} + \varrho b_{ik} \qquad (i, k = 1, 2, \ldots, n)$$

und die Elemente f_{ik} der äquivalenten Matrix F von der Form

$$f_{ik} = g_{ik} + \varrho h_{ik} \qquad (i, k = 1, 2, \ldots, n).$$

Die Größen a_{ik}, b_{ik}, g_{ik} und h_{ik} bedeuten beliebige Konstanten aus dem Rationalitätsbereich Ω. Man beweist, daß, wenn die

zwei Determinanten $|b_{ik}|$ und $|h_{ik}|$ nicht verschwinden, in der Gleichung $F = RDQ$ die Systeme R und Q so gewählt werden können, daß *nicht nur ihre Determinanten, sondern auch ihre einzelnen Elemente von dem Parameter ϱ unabhängig sind.*

Lehrsatz X: Stimmen die zwei bilinearen Formen $A + \varrho B$ und $G + \varrho H$, bei denen die Determinanten $|B|$ und $|H|$ nicht verschwinden, in den zusammengesetzten Elementarteilern (oder den Determinantenteilern oder den einfachen Elementarteilern) überein, so kann man sie durch Substitutionen ineinander transformieren, deren Koeffizienten vom Parameter ϱ unabhängig sind.

Um Satz X von der Voraussetzung des Nichtverschwindens der zwei Determinanten $|b_{ik}|$ und $|h_{ik}|$ zu befreien, empfiehlt es sich, zur homogenen Betrachtung überzugehen.

Sind ϱ_1 und ϱ_2 unbestimmte Größen, so heißt

$$\varrho_1 A + \varrho_2 B = \sum_{i=1}^{i=n} \sum_{k=1}^{k=n} (\varrho_1 a_{ik} + \varrho_2 b_{ik}) x_i y_k$$

eine *Schar bilinearer Formen* (Kronecker (1868), *Ges. Werke* **1**, 166). A und B heißen die *Grundformen der Schar* ($\varrho_1 = 1$, $\varrho_2 = \varrho$ liefert den früher behandelten unhomogenen Fall $A + \varrho B$). Wir setzen wieder $D = \varrho_1 A + \varrho_2 B$. Verschwindet die Determinante der Schar identisch, d. h. ist sie für jedes Wertepaar ϱ_1 und ϱ_2 Null, so heißt die Schar *singulär*, im anderen Fall *ordinär*. Eine singuläre Schar ist dadurch charakterisiert, daß der Rang l von D kleiner als n ist; bei einer ordinären Schar ist $l = n$.

Um noch engeren Anschluß an Weierstraß zu gewinnen, lassen wir Ω mit dem größten aller möglichen Rationalitätsbereiche zusammenfallen, d. h. Ω *bedeutet die Gesamtheit aller Zahlen.* Die Elemente a_{ik}, b_{ik}, f_{ik} und g_{ik} der Matrices A, B, F und G sollen also jetzt beliebige Konstanten sein.

Da in Weierstraß' klassischen Untersuchungen die einfachen Elementarteiler, mit denen er allein operiert, direkt eingeführt werden, soll dies auch noch geschehen: Ist l der Rang der Matrix $D = \varrho_1 A + \varrho_2 B$, so sei σ eine ganze Zahl, die $\leqq l$ ist. Weierstraß selbst behandelt nur den Fall $l = n$, d. h. die Determinante $|D|$ verschwindet nicht für alle Werte ϱ_1, ϱ_2. Da Ω der Bereich aller Zahlen ist, werden die irreduziblen Funktionen p lineare Funktionen von ϱ_1 und ϱ_2. Wir setzen $p = a\varrho_1 + b\varrho_2$ und verstehen unter h_σ den Exponenten der *höchsten* Potenz, mit dem die lineare, homogene ganze Funktion p

in allen Unterdeterminanten D_σ σ^{ten} Grades von D enthalten ist. Ist T_σ der größte gemeinsame Teiler sämtlicher Unterdeterminanten σ^{ten} Grades D_σ, so hat der Determinantenteiler T_σ die Funktion p genau in der h_σ^{ten} Potenz zum Faktor. Sämtliche Unterdeterminanten D_σ sind durch p^{h_σ} teilbar, einige von ihnen können auch eine höhere Potenz von p zum Faktor haben, wenigstens eine Unterdeterminante D_σ muß entsprechend der Definition der Zahl h_σ die Funktion p genau in der Potenz h_σ enthalten. Eine solche D_σ, die genau durch die h_σ^{te} Potenz von p teilbar ist, heißt nach Frobenius (*Über die Elementarteiler der Determinanten, Sitzungsb. d. Berl. Akad.* (1894), 32) eine *in bezug auf p reguläre* D_σ.

Die Zahlen h_σ sind ganze positive Zahlen oder Null. Für sie beweist man die Ungleichung $h_{\sigma+1} \geqq h_\sigma$, und, wenn $h_\sigma = 0$ ist, so ist $h_{\sigma-1} = h_{\sigma-2} = \cdots = 0$.

Wir setzen:

$$\begin{aligned} e_l &= h_l - h_{l-1}, \\ e_{l-1} &= h_{l-1} - h_{l-2}, \\ &\vdots \\ e_\sigma &= h_\sigma - h_{\sigma-1}, \\ &\vdots \\ e_1 &= h_1. \end{aligned}$$

Hieraus folgt

$$h_\sigma = e_\sigma + e_{\sigma-1} + \cdots + e_1.$$

In T_σ tritt folglich $p^{h_\sigma} = p^{e_\sigma} \cdot p^{e_{\sigma-1}} \cdot p^{e_{\sigma-2}} \cdots p^{e_1}$

$(e_\sigma \geqq e_{\sigma-1} \geqq e_{\sigma-2} \cdots \geqq e_1)$ (Cayleysche Ungleichung, siehe S. 104)

als Faktor auf. Jeder einzelne der mit einem von Null verschiedenen Exponenten versehenen Faktoren

$$p^{e_\sigma},\ p^{e_{\sigma-1}},\ p^{e_{\sigma-2}},\ \ldots,\ p^{e_1},$$

in die p^{h_σ} zerlegt wurde, ist ein *Weierstraßscher oder einfacher Elementarteiler*. Hiermit sind die einfachen Elementarteiler unter Umgehung der zusammengesetzten definiert. Die Basen der Weierstraßschen Elementarteiler findet man, indem man T_l als Produkt von linearen homogenen Funktionen der Variablen ϱ_1 und ϱ_2 und eines von diesen Veränderlichen unabhängigen konstanten Faktors darstellt.

Der Weierstraßsche Fundamentalsatz lautet:

Lehrsatz XI: Zwei Scharen bilinearer Formen, deren Determinanten nicht verschwinden, sind dann und nur dann äquivalent (d. h. durch von den Parametern unabhängige, lineare homogene Substitutionen mit nichtverschwindenden Determinanten gegenseitig ineinander transformierbar), wenn die Determinanten der beiden Scharen in ihren Elementarteilern übereinstimmen (Weierstraß (1868), *Ges. Werke* **2**, 21).

Statt die einfachen Elementarteiler zu verwenden, wie es Weierstraß tut, kann man sich auch im Satze XI als Kriterium für die Äquivalenz gleichwertig der Übereinstimmung in den zusammengesetzten Elementarteilern E oder in den Determinantenteilern T bedienen. Die letzten beiden sind als größte gemeinsame Teiler rational zu finden, während die Bestimmung der einfachen Elementarteiler der Schar $\varrho_1 A + \varrho_2 B$ die Zerfällung einer Gleichung in lineare Faktoren, also irrationale Operationen, erfordert. *Über die Äquivalenz zweier ordinärer Scharen bilinearer Formen kann also auf rationalem Wege entschieden werden; auch die Substitutionen, die eine Schar in die andere überführen, sind rational bestimmbar.*

Weierstraß geht bei seinem Beweisverfahren durch das Irrationale hindurch. Er transformiert die Formenschar auf eine Normalform, die von den einfachen Elementarteilern abhängt. *Man kann stets eine ordinäre Schar bilinearer Form mit vorgegebenen Elementarteilern konstruieren.* Dies ergibt sich auch aus der weiter unten angegebenen Normalform einer linearen homogenen Substitution. Frobenius (*Journ. f. Math.* **86**, 146 (1879)) hat zuerst den Weierstraßschen Fundamentalsatz XI nur mit Hülfe durchaus rationaler Operationen bewiesen (vgl. auch Landsberg, *Journ. f. Math.* **116**, 331 (1896)). Über die an den Weierstraßschen Fundamentalsatz anknüpfende Literatur vgl. man Muth, *Theorie und Anwendung der Elementarteiler*, S. 60. Wir heben hier nur die Arbeit von Stickelberger, *Journ. f. Math.* **86**, 20 (1879), hervor.

Aus dem Weierstraßschen Fundamentalsatz XI folgt:
Damit es möglich sei, die Formenschar:

$$\sum_{i=1}^{i=n} \sum_{k=1}^{k=n} (\varrho_1 a_{ik} + \varrho_2 b_{ik}) x_i y_k$$

von nichtverschwindender Determinante durch eine lineare homogene Substitution für $x_1, x_2, \ldots, x_n$ von nichtverschwindender Determinante und eine andere für $y_1, y_2, \ldots, y_n$ von ebenfalls nichtverschwindender Determinante in die Normalform:

$$\varrho_1(a_1 x_1' y_1' + a_2 x_2' y_2' + \cdots + a_n x_n' y_n') +$$
$$\varrho_2(b_1 x_1' y_1' + b_2 x_2' y_2' + \cdots + b_n x_n' y_n')$$

zu transformieren, ist notwendig und hinreichend, daß die Determinante $|\varrho_1 a_{ik} + \varrho_2 b_{ik}|$ *nur Elementarteiler mit den Exponenten 1 besitzt, d. h. jeder Teiler* $a\varrho_1 + b\varrho_2$ *der Determinante* $|\varrho_1 a_{ik} + \varrho_2 b_{ik}|$, *der in ihr* $s > 1$ *mal enthalten ist, auch Teiler aller ihrer* $n - 1^{ten}$, $n - 2^{ten}$, *bis* $n - s + 1^{ten}$ *Unterdeterminanten ist* (Weierstraß, a. a. O., S. 37).

Die Äquivalenz zweier *singulärer* Scharen bilinearer Formen hat Kronecker (*Algebraische Reduktion der Scharen bilinearer Formen, Sitzungsb. d. Berl. Akad.* (1890), 1225) erledigt. Wir geben das Resultat ohne Benützung alles Voraufgegangenen, indem wir hierbei auch noch die Äquivalenz ordinärer Scharen mittels der Determinantenteiler besonders hervorheben:

Sei

$$D = \sum_{i=1}^{i=n} \sum_{k=1}^{k=n} (\varrho_1 a_{ik} + \varrho_2 b_{ik}) x_i y_k$$

eine Schar bilinearer Formen (bei der auch infolge verschwindender Koeffizienten die Anzahl beider Reihen von Veränderlichen verschieden sein kann). T_σ *bedeutet den größten gemeinsamen Teiler aller Unterdeterminanten* σ^{ten} *Grades* ($\sigma = 1, 2, \ldots$). *Verschwindet die Determinante der Schar nicht, so sind alle diejenigen Scharen und auch nur diejenigen mit D äquivalent, die mit D in den so ermittelten* T_σ *übereinstimmen* (andere Formulierung des Weierstraßschen Satzes XI). *Verschwindet die Determinante von D für alle Werte von* ϱ_1 *und* ϱ_2, *so ist die obige Bedingung nicht ausreichend. Sei l der Rang der Determinante von D, so bestehen zwischen den partiellen Ableitungen von* $\varrho_1 A + \varrho_2 B$ *nach den n Variablen* x_i *und den n Variablen* y_i *genau zweimal* $n - l$ *linear unabhängige Relationen, deren Koeffizienten homogene Funktionen von* ϱ_1 *und* ϱ_2 *sind. Denkt man sich ein vollständiges System von* $2n - 2l$ *solchen Relationen von möglichst niedrigem Grade in* ϱ_1, ϱ_2 *gewählt, so sind alle außer in den Größen* T_σ *noch in diesen Minimalgradzahlen*[1]) *mit D übereinstimmenden Scharen und auch nur diese mit unserer Schar D äquivalent* (Kronecker).

Charakteristische Kriterien für die Äquivalenz zweier singulärer Scharen sind natürlich auch die Übereinstimmung in den

[1]) Die Bezeichnung stammt von Muth, *Theorie der Elementarteiler*, S. 108.

Minimalgradzahlen und den einfachen Elementarteilern oder in den Minimalgradzahlen und den zusammengesetzten Elementarteilern. Sowohl das oben ausführlich gegebene Theorem als auch die letzte Bedingung gestatten über die Äquivalenz zweier singulärer Scharen *rational*, nämlich durch Aufsuchen größter gemeinsamer Teiler, eine Entscheidung zu fällen. Auch die *überführenden Substitutionen* sind *rational* zu finden. — Für gegebenes n kann man stets eine singuläre Formenschar von n Variablenpaaren konstruieren, welche die vorgeschriebenen Minimalgradzahlen und Elementarteiler besitzt.

Gehören die Koeffizienten der Grundformen A und B der Schar $\varrho_1 A + \varrho_2 B$ einem Rationalitätsbereiche Ω an und sind die Koeffizienten der Grundformen der äquivalenten Schar ebenfalls nur Zahlen aus dem nämlichen Bereich Ω, so kann man auch die Koeffizienten der die zwei Scharen ineinander überführenden Substitutionen aus dem Rationalitätsbereich Ω wählen. Will man nur im Rationalitätsbereich Ω operieren, so hat man als Basen einfacher Elementarteiler der Matrix

$$\| \varrho_1 a_{ik} + \varrho_2 b_{ik} \| \qquad (i, k = 1, 2, \ldots, n)$$

anstatt der linearen homogenen Funktionen $a\varrho_1 + b\varrho_2$ solche homogene Funktionen der zwei Variablen ϱ_1, ϱ_2 mit Koeffizienten aus Ω zu verwenden, die bezüglich des Rationalitätsbereiches Ω irreduzibel sind.

Die Untersuchung einer ganzzahligen Matrix, also die rein zahlentheoretische Behandlung unseres Gegenstandes, mit der unsere Darstellung anhebt, verdankt man Henry J. St. Smith (1861). Weierstraß' Ausgangspunkt und ausschließlicher Gegenstand der Untersuchung ist eine ordinäre Schar bilinearer Formen, für die er als charakteristisches Kennzeichen der Äquivalenz die einfachen Elementarteiler ersonnen hat. Inwieweit die Priorität der Entdeckung der Elementarteiler für Sylvester (1851) (*Coll. math. papers* **1**, 219) in Anspruch genommen werden kann, vgl. man bei Noether, *Math. Ann.* **50**, 137. Durch Sylvester angeregt, fand Cayley die oben erwähnte Ungleichung. Die Brücke zwischen den Untersuchungen von Smith und Weierstraß hat Frobenius geschlagen (*Journ. f. Math.* **86**, 146 (1879), 88, 96 (1880), *Sitzungsb. d. Berl. Akad.* (1894), 31). Er hat die große Allgemeinheit des Begriffes „Elementarteiler" gezeigt; ihm verdankt man auch den für die ganze Theorie fundamentalen Satz IV, der für Matrices mit Elementen, die ganze Zahlen oder ganze Funktionen einer oder mehrerer Variablen

mit beliebigen konstanten Koeffizienten oder Größen aus einem Rationalitätsbereich Ω sind, gilt. Die Untersuchung der singulären Scharen bilinearer Formen hat Kronecker während mehr als zwanzig Jahren beschäftigt, bis er zu dem oben angegebenen abschließenden Resultate gelangte.

Der Begriff „Elementarteiler" ist im voraufgehenden immer nur bei ganzen Zahlen oder bei ganzen Funktionen verwendet worden. Er läßt sich auch auf gebrochene Größen ausdehnen. Vgl. Hensel, *Journ. f. Math.* **115**, 259 (1895), sowie seine Bearbeitung der *Kroneckerschen Vorlesungen über Determinanten*, Vorl. 21.

Von Lehrbüchern sind vor allem Muth, *Theorie der Elementarteiler*, dann die eben zitierte *Vorlesung* von Kronecker und für die zahlentheoretischen Fragen Bachmann, *Die Arithmetik der quadratischen Formen*, Leipzig 1898, S. 288, zu nennen.

Am Schlusse des § 6 wiesen wir darauf hin, daß zwei beliebige bilineare Formen gleichen Ranges nicht stets kogredient ineinander transformierbar sind. Für die kogrediente Transformation zweier bilinearer Formen gilt der Satz von Kronecker (*Monatsb. d. Berl. Akad.* (1874), *Ges. Werke* **1**, 423):

Damit zwei einzelne bilineare Formen A und B durch kogrediente Substitutionen ineinander transformiert werden können, ist notwendig und hinreichend, daß die zwei Scharen $\varrho_1 A + \varrho_2 A'$ und $\varrho_1 B + \varrho_2 B'$ äquivalent sind; A' und B' sind die zu A und B transponierten Formen (vgl. den einfachen Beweis von Frobenius, *Sitzungsb. d. Berl. Akad.* (1896), 14).

Zwei beliebige bilineare Formen, lineare homogene Substitutionen oder Matrices *sind nicht ähnlich* (vgl. S. 84). *Über die Ähnlichkeit zweier Systeme* geben wir noch folgendes:

Zwei Substitutionen sind dann und nur dann ähnlich, wenn ihre charakteristischen Funktionen die nämlichen Elementarteiler haben (Weierstraß, *Ges. Werke* **2**, 21 u. 22).

Die charakteristische Funktion $|\varrho E_i - C_i|$, wobei E_i die Einheitsmatrix e_i^{ten} Grades und C_i die Matrix

$$\left\| \begin{array}{ccccccc} c_i & 0 & 0 & \dots\dots & 0 \\ 1 & c_i & 0 & \dots\dots & 0 \\ 0 & 1 & c_i & \dots\dots & 0 \\ 0 & 0 & 1 & c_i \dots\dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 \dots\dots 1 & c_i \end{array} \right\| \quad (e_i \text{ Zeilen und } e_i \text{ Kolonnen})$$

oder die lineare homogene Substitution

$$y_{i1} = c_i x_{i1},\ y_{i2} = x_{i1} + c_i x_{i2},\ y_{i3} = x_{i2} + c_i x_{i3},\ \ldots,$$
$$y_{i e_i} = x_{i e_{i-1}} + c_i x_{i e_i}$$

ist, hat den *einzigen Elementarteiler* $(\varrho - c_i)^{e_i}$; die Determinante $|C_i|$ kann hierbei von Null verschieden sein oder auch verschwinden.

Jede lineare homogene Substitution oder Matrix C, deren charakteristische Funktion $|\varrho E - C|$ *die Elementarteiler* $(\varrho - c_1)^{e_1}$, $(\varrho - c_2)^{e_2}, \ldots, (\varrho - c_i)^{e_i}$ *besitzt, ist der zerlegbaren Substitution oder Matrix* $C_1 + C_2 + C_3 + \cdots + C_\tau$ (vgl. S. 92) *ähnlich;* C_i $(i = 1, 2, \ldots, \tau)$ *hat die obige Bedeutung.* Diese *Normalform* wird nur durch die Elementarteiler von $|\varrho E - C|$ bestimmt. Über die Verwertung dieser Normalform in der Theorie der linearen homogenen Differentialgleichungen vgl. etwa C. Jordan, *Cours d'analyse,* Paris 1896, **3**, 173, J. Horn, *Gewöhnliche Differentialgleichungen, Samml. Schubert* **50**, Leipzig 1905, S. 68, L. Schlesinger, *Vorl. über die Theorie der linearen Differentialgleichungen,* Leipzig 1908, 8. Vorl.

Auch wenn die Substitution C Koeffizienten aus einem endlichen Körper mit p^s Größen, dem sogenannten $GF[p^s]$, hat (vgl. Kap. III, § 1), kann man für C Normalformen von der Art der obigen aufstellen. Vgl. C. Jordan, *Traité des substitutions et des équations algébriques,* Paris 1870, S. 114—126, Frobenius, *Journ. f. Math.* **86**, 207 (1879), L. E. Dickson, *Linear groups, Teubners Samml.* **6**, Leipzig 1901, S. 221. Auch hier hängen die Normalformen mit dem Begriff „Elementarteiler" zusammen; denn dieser ist auch noch anwendbar, wenn der Rationalitätsbereich Ω nicht mehr unendlich viele, sondern nur eine endliche Anzahl von Konstanten enthält, nämlich das $GF[p^s]$ ist.

Ist die Determinante $|C|$ einer linearen homogenen Substitution nicht gleich Null, so sind für $k = \pm 1, \pm 2, \pm 3, \ldots$

$$(\varrho - c_1^k)^{e_1},\ (\varrho - c_2^k)^{e_2},\ \ldots,\ (\varrho - c_\tau^k)^{e_\tau}$$

die Elementarteiler der charakteristischen Funktion der Potenz C^k. Da die Elementarteilerexponenten $e_1, e_2, \ldots, e_\tau$ für alle Potenzen von C die nämlichen sind, lassen sich diese Exponenten auch durch das Studium der *von C und seinen positiven wie negativen Potenzen erzeugten zyklischen Gruppe* gewinnen. J. Wirth, *Freiburger Diss.* (1906).

§ 9. Die quadratischen und Hermiteschen Formen. Die alternierenden bilinearen Formen.

Eine *symmetrische bilineare Form*

$$A = \sum_{i=1}^{i=n} \sum_{k=1}^{k=n} a_{ik} x_i y_k \qquad (a_{ik} = a_{ki})$$

wird zur *quadratischen Form:*

$$f = \sum_{i=1}^{i=n} \sum_{k=1}^{k=n} a_{ik} x_i x_k \qquad (a_{ik} = a_{ki}),$$

wenn man die zwei Variablenreihen x_i und y_i gleichsetzt.

Die zu der quadratischen Form

$$f = \sum_{i=1}^{i=n} \sum_{k=1}^{k=n} a_{ik} x_i x_k \qquad (a_{ik} = a_{ki})$$

zugehörige symmetrische bilineare Form:

$$A = \frac{1}{2}\left(y_1 \frac{\partial f}{\partial x_1} + y_2 \frac{\partial f}{\partial x_2} + \cdots + y_n \frac{\partial f}{\partial x_n}\right)$$

heißt die *Polarform von f.* Zu der quadratischen Form f gehört die symmetrische Matrix

$$A = \| a_{ik} \| \qquad (i, k = 1, 2, \ldots, n)$$

mit den Elementen $a_{ik} = a_{ki}$.

Die Determinante von A:

$$|A| = |a_{ik}| \qquad (i, k = 1, 2, \ldots, n)$$

heißt die *Determinante* oder die *Diskriminante* der quadratischen Form; $2|A|$ ist die *Hessesche Determinante von f* (vgl. § 15).

Ist $A \neq 0$, so ist f eine *ordinäre,* sonst eine *singuläre quadratische Form.*

Die quadratische Form:

$$F = \sum_{i=1}^{i=n} \sum_{k=1}^{k=n} A_{ik} u_i u_k = -\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} & u_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & u_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & u_n \\ u_1 & u_2 & \cdots & u_n & 0 \end{vmatrix},$$

bei der A_{ik} die algebraischen Adjungierten der Elemente a_{ik} in

der Determinante $|A|$ sind, heißt die zu f *adjungierte quadratische Form* (Gauß, *Disquisitiones arithmeticae,* Artikel 267, *Ges. Werke* **1**, 301, hat für $-F$ die Bezeichnung „forma adjuncta").

Ist $|A| = 0$, *so ist die adjungierte Form* F *das volle Quadrat einer linearen Form. Ist der Rang der Determinante* $|A|$ *kleiner als* $n-1$, *so verschwindet* F *identisch.*

Für jede quadratische Form f *besteht die Relation:*

$$-|A|^{n-2} \cdot f = \begin{vmatrix} A_{11} & A_{12} & \dots & A_{1n} & x_1 \\ A_{21} & A_{22} & \dots & A_{2n} & x_2 \\ \vdots & \vdots & & \vdots & \vdots \\ A_{n1} & A_{n2} & \dots & A_{nn} & x_n \\ x_1 & x_2 & \dots & x_n & 0 \end{vmatrix}.$$

Führt die lineare homogene Substitution P*:*

$$x_i = p_{i1}x_1' + p_{i2}x_2' + \cdots + p_{in}x_n' \qquad (i = 1, 2, \dots, n)$$

die quadratische Form $\sum_{i=1}^{i=n}\sum_{k=1}^{k=n} a_{ik}x_ix_k$ *in die quadratische Form* $\sum_{i=1}^{i=n}\sum_{k=1}^{k=n} b_{ik}x_i'x_k'$ *über, so gehen die Polarformen der zwei quadratischen Formen durch* P *und eine zu* P *kogrediente Substitution ineinander über. Es besteht die symbolische Gleichung* $B = P'AP$, *wobei* P' *die zu* P *transponierte Matrix ist.*

Hieraus folgt: *Die Determinante* $|B|$ *ist gleich* $|P|^2 \cdot |A|$.

Ferner ergibt sich:

Zwei quadratische Formen können dann und nur dann durch eine lineare homogene Substitution von nichtverschwindender Determinante ineinander übergeführt werden, wenn sie gleichen Rang l *besitzen. Insbesondere kann jede quadratische Form vom Range* l *in die Normalform*

$$a_1x_1'^2 + a_2x_2'^2 + \cdots + a_lx_l'^2,$$

wobei $a_1, a_2, \dots, a_l$ *lauter nichtverschwindende, sonst beliebige Größen sind, transformiert werden.*

Das Problem der Reduktion einer quadratischen Form auf eine *lineare Kombination von Quadraten* ist überall in Geometrie und Mechanik, wo quadratische Formen auftreten, von größter Bedeutung. Ein allgemein gültiges, besonders einfaches Verfahren, *jede* beliebige quadratische Form in eine Summe von

Quadraten umzuwandeln, besteht in einer sukzessiven Abtrennung von 1 oder 2 Quadraten (Baltzer, *Determinanten*, 5. Aufl., S. 175, J. A. Serret, *Algebra*, deutsche Ausg. von Wertheim, Leipzig 1878, S. 350). Diese Methode knüpft an Lagrange (*Œuvres* **1**, 7, *Recherches sur la méthode de maximis et minimis* (1759)) an und wird daher auch als *„Lagrangesche Transformation"* (Gundelfinger, *Journ. f. Math.* **91**, 221 (1881), ebenda **127**, 86 (1904)) bezeichnet.

Man kann die Variablen einer beliebigen quadratischen Form

$$\sum_{i=1}^{i=n}\sum_{k=1}^{k=n} a_{ik}x_ix_k$$

vom Range l stets in einer solchen Reihenfolge mit

$$x_{\sigma_1},\ x_{\sigma_2},\ \ldots,\ x_{\sigma_l}$$

bezeichnen, daß

$$\Gamma_l = \sum \pm a_{\sigma_1\sigma_1}a_{\sigma_2\sigma_2}\ldots a_{\sigma_l\sigma_l}$$

von Null verschieden ist und in der Kette der Hauptunterdeterminanten:

$$\Gamma_l,\ \Gamma_{l-1} = \frac{\partial\Gamma_l}{\partial a_{\sigma_l\sigma_l}},\ \Gamma_{l-2} = \frac{\partial^2\Gamma_l}{\partial a_{\sigma_l\sigma_l}\partial a_{\sigma_{l-1}\sigma_{l-1}}},$$

$$\ldots,\ \Gamma_2 = a_{\sigma_1\sigma_1}a_{\sigma_2\sigma_2} - a^2_{\sigma_1\sigma_2},\ \Gamma_1 = a_{\sigma_1\sigma_1},\ \Gamma_0 = 1$$

nie zwei unmittelbar aufeinanderfolgende gleichzeitig verschwinden (vgl. die beim letzten Satz über symmetrische Determinanten auf S. 63 angeführte Literatur).

Jede quadratische Form

$$f = \sum_{i=1}^{i=n}\sum_{k=1}^{k=n} a_{ik}x_ix_k$$

vom Range l kann, falls ihre Variablen solcher Anordnung fähig sind, daß keine der Größen $\Gamma_1, \Gamma_2, \ldots, \Gamma_l$ *verschwindet, in der Form:*

$$\frac{X_1^2}{\Gamma_1} + \frac{X_2^2}{\Gamma_1\Gamma_2} + \frac{X_3^2}{\Gamma_2\Gamma_3} + \cdots + \frac{X_l^2}{\Gamma_{l-1}\Gamma_l}$$

dargestellt werden. Hierbei ist

$$X_1 = \frac{1}{2}\frac{\partial f}{\partial x_{\sigma_1}},$$

$$X_\alpha = \begin{vmatrix} a_{\sigma_1\sigma_1} & \cdots & a_{\sigma_1\sigma_{\alpha-1}} & \frac{1}{2}\frac{\partial f}{\partial x_{\sigma_1}} \\ a_{\sigma_2\sigma_1} & \cdots & a_{\sigma_2\sigma_{\alpha-1}} & \frac{1}{2}\frac{\partial f}{\partial x_{\sigma_2}} \\ \vdots & & \vdots & \vdots \\ a_{\sigma_\alpha\sigma_1} & \cdots & a_{\sigma_\alpha\sigma_{\alpha-1}} & \frac{1}{2}\frac{\partial f}{\partial x_{\sigma_\alpha}} \end{vmatrix} \qquad (\alpha = 2, 3, \ldots, l).$$

Die Funktion X_α enthält keine der $\alpha - 1$ Variablen $x_{\sigma_1}, x_{\sigma_2}, \ldots, x_{\sigma_{\alpha-1}}$. Diese Transformation bezeichnet man als Jacobische Transformation der quadratischen Form (Jacobi, *Journ. f. Math.* **53**, 265 (1857), *Ges. Werke* **3**, 583, hat diese Transformation allgemein für bilineare Formen aufgestellt).

Im Gegensatz zu der an Lagranges Namen anknüpfenden Methode, die ausnahmslos bei jeder quadratischen Form angewendet werden kann, involviert die Jacobische Transformation eine Annahme. Sie setzt voraus, die Variablen der quadratischen Form f lassen sich wenigstens auf eine Weise so anordnen, daß jede der Determinanten $\Gamma_l, \Gamma_{l-1}, \ldots, \Gamma_1, \Gamma_0$ von Null verschieden ist. Die Jacobische Transformation versagt z. B. schon bei quadratischen Formen f, deren Hauptelemente $a_{11}, a_{22}, \ldots, a_{nn}$ sämtlich verschwinden. Ist bei einer quadratischen Form die Jacobische Transformation anwendbar, so ist sie im Effekt mit Lagranges Reduktionsmethode identisch. Auch Gauß (*Ges. Werke* **4**, 37, **6**, 20, 7a, 248) hat sich bei seinen astronomischen Untersuchungen der Jacobischen Transformation für quadratische Formen bedient. Ihre Ausdehnung auf bilineare Formen wie die explizite Darstellung der X_α durch die x_i ist Jacobis bleibendes Verdienst.

Eine quadratische Form kann auf unendlich viele Arten in eine Summe von Quadraten transformiert werden. Hat die gegebene Form *reelle* Koeffizienten und soll die auf die Variablen auszuführende Substitution *gleichfalls reell* sein, so gilt das von Sylvester (*Phil. Mag.* (1852), 142, *Phil. Trans.* (1853), 481, *Coll. math. papers* **1**, 380 u. 511) gefundene und von ihm sogenannte *Trägheitsgesetz der reellen quadratischen Formen: Auf welche Weise auch immer eine beliebige quadratische Form mit reellen Koeffizienten durch eine reelle lineare homogene Substitution von nichtverschwindender Determinante in die Normalform*

$$x_1'^2 + x_2'^2 + \cdots + x_h'^2 - x_{h+1}'^2 - x_{h+2}'^2 \cdots - x_l'^2$$

transformiert wird, stets haben außer dem Range l die Anzahl h der Quadrate mit positiven Vorzeichen und folglich auch die Anzahl $l - h$ der Quadrate mit negativen Vorzeichen dieselben unveränderlichen Werte.

Das Trägheitsgesetz war bereits Riemann aus Gauß' Vorlesungen über die Methode der kleinsten Quadrate, die Riemann wahrscheinlich 1846/47 hörte, bekannt (Riemanns *Ges. Werke,* Nachträge, herausg. von M. Noether u. W. Wirtinger, Leipzig 1902, S. 59). Auch Jacobi war seit 1847, also vor Sylvesters Publikation, auf das Trägheitsgesetz gekommen, vgl. den posthumen Aufsatz von Jacobi, *Journ. f. Math.* **53**, 275 (1857), *Ges. Werke* **3**, 591, sowie den unmittelbar voraufgehend abgedruckten Aufsatz von Hermite (*Œuvres* **1**, 429) und die an diese Aufsätze anknüpfenden Bemerkungen von Borchardt (*Ges. Werke*, S. 469).

Die Zahl h kann durch irgendeine reelle Transformation der reellen quadratischen Form in eine Summe von Quadraten gefunden werden. Eine Bestimmung der Zahl h kann auf folgende Weise geschehen: Man ordne die Indices der reellen quadratischen Form

$$\sum_{i=1}^{i=n} \sum_{k=1}^{k=n} a_{ik} x_i x_k$$

derartig, daß nie zwei aufeinanderfolgende der Determinanten

$$\Gamma_l, \Gamma_{l-1}, \Gamma_{l-2}, \ldots, \Gamma_2, \Gamma_1, \Gamma_0$$

gleichzeitig verschwinden (vgl. oben). Ist $\Gamma_k = 0$, so haben Γ_{k-1} und Γ_{k+1} entgegengesetzte Vorzeichen. Die reelle quadratische Form

$$\sum_{i=1}^{i=n} \sum_{k=1}^{k=n} a_{ik} x_i x_k$$

enthält dann bei einer reellen Transformation in eine Summe von Quadraten genau ebensoviele negative Quadrate ($l - h$), wie die Reihe der angegebenen Determinanten Vorzeichenwechsel aufweist; etwa vereinzelt auftretende Nullen sind fortzulassen (Gundelfinger, Zusätze zur dritten Auflage von Hesses *Vorl. üb. analy Geometrie d. Raumes,* Leipzig 1876, S. 460, *Journ. f. Math.* **91**, 225 (1881)).

An Stelle der Größen Γ kann man eine allgemeinere, von Darboux (*Journ. de math.* (2) **19**, 352 (1874), Gundelfinger, *Hesses Vorl.*, S. 459) stammende Kette von Determinanten zur

Bestimmung der Zahl h verwenden; in dieser Kette sind die Γ als Spezialfall enthalten. Auf S. 126 findet man noch ein Kriterium von Sylvester zur Bestimmung von h.

Die Zahl $2h - l$ heißt nach Frobenius, *Über das Trägheitsgesetz der quadratischen Formen, Journ. f. Math.* **114**, 187 (1895) die *Signatur* der reellen quadratischen Form. Frobenius gibt ein Verfahren zur Bestimmung der Signatur einer quadratischen Form, falls mehrere aufeinanderfolgende Determinanten Γ verschwinden. Diese Methode schränkt die vorherige Umordnung der Indices möglichst ein.

Die kleinere der zwei Zahlen h und $l - h$ heißt die *Charakteristik* der reellen quadratischen Form (A. Loewy, *Journ. f. Math.* **122**, 53 (1900)). Über Untersuchungen, die mit der Charakteristik der reellen quadratischen Formen zusammenhängen, vgl. S. 127 u. 136.

Ist die Charakteristik einer reellen quadratischen Form Null, so heißt eine Form von nichtverschwindender Determinante definit, von verschwindender Determinante semidefinit. Alle anderen reellen quadratischen Formen heißen *indefinit* (Gauß, *Disquisitiones arithmeticae*, Art. 271).

Eine definite quadratische Form nimmt nur Werte eines und desselben Vorzeichens an, welche reellen Werte man auch den Variablen beilegen mag; sie verschwindet nur, wenn man alle Variablen Null setzt. Eine semidefinite Form nimmt ebenfalls nur Werte eines und desselben Vorzeichens an; sie verschwindet aber auch, ohne daß man alle Variablen Null zu setzen braucht.

Ein wichtiges Problem ist es, zu entscheiden, *ob eine Schar quadratischer Formen mit beliebigen Koeffizienten*:

$$\sum_{i=1}^{i=n}\sum_{k=1}^{k=n}(\varrho_1 a_{ik} + \varrho_2 b_{ik})\,x_i x_k$$

durch eine lineare homogene Substitution P:

$$x_i = p_{i1}x_1' + p_{i2}x_2' + \cdots + p_{in}x_n'$$

von nichtverschwindender Determinante in eine andere Schar

$$\sum_{i=1}^{i=n}\sum_{k=1}^{k=n}(\varrho_1 g_{ik} + \varrho_2 h_{ik})\,x_i' x_k'$$

transformiert werden kann. Ist dies der Fall, so führt die Substitution P die Scharen symmetrischer bilinearer Formen:

$$\sum_{i=1}^{i=n}\sum_{k=1}^{k=n}(\varrho_1 a_{ik}+\varrho_2 b_{ik})x_i y_k$$

und

$$\sum_{i=1}^{i=n}\sum_{k=1}^{k=n}(\varrho_1 g_{ik}+\varrho_2 h_{ik})x_i' y_k'$$

als Polarformen kogredient ineinander über. Es gilt die symbolische Gleichung:

$$P'(\varrho_1 A+\varrho_2 B)P=\varrho_1 G+\varrho_2 H.$$

Mithin ergibt sich:

Alle für die Transformation der Scharen bilinearer Formen (vgl. S. 113 u. 114) *gefundenen Bedingungen sind auch für die Transformation der quadratischen Formen notwendig. Daß sie auch ausreichend sind, ist ein bemerkenswertes Resultat, das für Scharen von nichtverschwindender Determinante Weierstraß in seiner oben* (§ 8) *mehrfach zitierten Abhandlung aus dem Jahre 1868, für solche von verschwindender Determinante Kronecker* (abschließende Resultate, *Sitzungsb. d. Berl. Akad.* (1890), 1375, (1891), 9 u. 33) *gewonnen hat.* Man vgl. vor allem Frobenius' fundamentale Arbeit *„Über die kogredienten Transformationen der bilinearen Formen", Sitzungsb. d. Berl. Akad.* (1896), 7, dort wird der eigentliche Grund dieser merkwürdigen Erscheinung aufgedeckt.

Es gilt also der Satz:

Zwei Scharen quadratischer Formen, deren Determinanten nicht verschwinden, lassen sich dann und nur dann ineinander transformieren, wenn ihre Determinanten in den Elementarteilern übereinstimmen, bei solchen mit verschwindenden Determinanten ist außerdem noch die Gleichheit der Minimalgradzahlen erforderlich (vgl. hierzu die Darstellung bei Muth, *Theorie der Elementarteiler,* S. 125).

Damit sich die Schar quadratischer Formen

$$\sum_{i=1}^{i=n}\sum_{k=1}^{k=n}(\varrho_1 a_{ik}+\varrho_2 b_{ik})x_i x_k$$

von nichtverschwindender Determinante durch eine lineare homogene Substitution von nichtverschwindender Determinante in die Gestalt:

$$\varrho_1(g_1 x_1'^2+g_2 x_2'^2+\cdots+g_n x_n'^2)+$$
$$+\varrho_2(h_1 x_1'^2+h_2 x_2'^2+\cdots+h_n x_n'^2)$$

bringen läßt, wobei g_i und h_i nie gleichzeitig verschwinden $(i = 1, 2, \ldots, n)$, *ist notwendig und hinreichend, daß die Determinante* $|\varrho_1 a_{ik} + \varrho_2 b_{ik}|$ *nur Elementarteiler mit den Exponenten 1 besitzt* (Weierstraß, *Ges. Werke* **2**, 41—42).

Ist in dem eben besprochenen Fall die Determinante von B ungleich Null, so braucht man nur

$$z_1 = \sqrt{h_1}\, x_1', \; z_2 = \sqrt{h_2}\, x_2', \; \ldots, \; z_n = \sqrt{h_n}\, x_n'$$

zu setzen, dann erhält man die quadratische Form:

$$W = \frac{g_1}{h_1} z_1^2 + \frac{g_2}{h_2} z_2^2 + \cdots + \frac{g_n}{h_n} z_n^2 = w_1 z_1^2 + w_2 z_2^2 + \cdots + w_n z_n^2$$

und die Einheitsform:

$$E = z_1^2 + z_2^2 + \cdots + z_n^2.$$

Die Größen

$$w_i = \frac{g_i}{h_i} \qquad (i = 1, 2, \ldots, n)$$

sind die Wurzeln der Gleichung

$$|\varrho b_{ik} - a_{ik}| = 0 \qquad (i, k = 1, 2, \ldots, n),$$

wie aus der symbolischen Gleichung $P'(\varrho B - A)P = \varrho E - W$ folgt.

Ist die gegebene quadratische Form B die Einheitsform E, so folgt:

Man kann die Schar quadratischer Formen $\varrho_1 A + \varrho_2 E$ dann und nur dann in $\varrho_1 W + \varrho_2 E$ transformieren, falls alle Elementarteiler der charakteristischen Funktion $|\varrho E - A|$ die Exponenten 1 haben. Die sich ergebende symbolische Gleichung $P'P = E$ besagt (vgl. S. 130), daß die überführende Transformation P orthogonal ist. Mithin hat man das Resultat:

Eine beliebige quadratische Form kann dann und nur dann orthogonal in eine Summe von Quadraten transformiert werden, falls alle Elementarteiler der charakteristischen Funktion den Exponenten 1 besitzen.

Die orthogonale Transformation einer quadratischen Form A in eine Quadratsumme haben für den sogenannten allgemeinen Fall, in dem die charakteristische Gleichung von A nur verschiedene Wurzeln besitzt, bereits Cauchy (*Exercices de math.* **4**, (1829), *Œuvres* (2) **9**, 194) und Jacobi (*Journ. f. Math.* **12**, 1 (1834), *Ges. Werke* **3**, 191) zum Gegenstand eingehender Untersuchungen gemacht.

Die charakteristische Funktion einer *reellen* quadratischen Form (vgl. die Angaben über die Säkulargleichung auf S. 65) besitzt ausschließlich reelle Wurzeln und Elementarteiler mit den Exponenten 1. Hieraus folgt:

Jede reelle quadratische Form A kann reell orthogonal in die reelle quadratische Form $w_1 z_1^2 + w_2 z_2^2 + \cdots + w_n z_n^2$ *transformiert werden; hierbei sind* $w_1, w_2, \ldots, w_n$ *die reellen Wurzeln der Säkulargleichung* $|A - \varrho E| = 0$.

Hat die charakteristische Funktion einer reellen quadratischen Form A von n Variablen die l_1*fache Wurzel Null, so ist der Rang l von A gleich* $n - l_1$. *Nach dem Trägheitsgesetz ist die für A charakteristische Zahl h der positiven Quadrate gleich der Anzahl der positiven Wurzeln der Säkulargleichung, die Anzahl der negativen Quadrate gleich der Anzahl der negativen Wurzeln.*

Eine Gleichung mit reellen Koeffizienten und lauter reellen Wurzeln besitzt genau soviel positive Wurzeln wie die Gleichung Zeichenwechsel hat (vgl. Kap. Algebra). Nun ist

$$|A - \varrho E| = (-1)^n \cdot (\varrho^n - T_1 \varrho^{n-1} + T_2 \varrho^{n-2} \cdots (-1)^n T_n).$$

Hierbei ist T_{n-s} die Summe der Hauptunterdeterminanten

$$\frac{\partial^s |A|}{\partial a_{i_1 i_1} \partial a_{i_2 i_2} \cdots \partial a_{i_s i_s}},$$

wobei $i_1, i_2, \ldots, i_s$ jede Kombination der Zahlen $1, 2, \ldots, n$ zu s bedeutet. Es ist also $T_n = |A|$, $T_{n-1} = \sum_{i=1}^{i=n} \frac{\partial |A|}{\partial a_{ii}}, \ldots,$ $T_1 = a_{11} + a_{22} + \cdots + a_{nn}$.

Die Zahl h der positiven Quadrate ist gleich der Anzahl der Vorzeichenfolgen der Determinantensummen $T_0 = 1, T_1, T_2, \ldots, T_n$[1]) (Satz von Sylvester, *Philos. Magazine* (1852), *Coll. math. papers* **1**, 380, vgl. Netto, *Vorl. über Algebra,* Leipzig 1896, S. 222, H. Weber, *Algebra* **1**, 311, Heffter, *Journ. f. Math.* **126**, 83 (1903), Gundelfinger, ebenda **127**, 85 (1904)).

Für beliebige Scharen quadratischer Formen von nichtverschwindender oder von verschwindender Determinante läßt sich stets eine Reduktion auf *kanonische* oder *Normalformen*

---

[1]) Abgesehen von den letzten Größen können in der angegebenen Kette nur isolierte T den Wert Null annehmen; die am Ende verschwindenden Glieder sind fortzulassen, die intermediär verschwindenden Größen sind nach Belieben positiv oder negativ zu zählen.

ausführen, deren Natur durch die Elementarteiler bestimmt wird. Zu vgl. Muth, *Theorie der Elementarteiler*, S. 124 u. 133.

Bei einer Schar reeller quadratischer Formen kann man aus den Charakteristiken (vgl. S. 123) der in der Schar enthaltenen Formen auf die Elementarteiler schließen:

Ist B eine reelle quadratische Form von nichtverschwindender Determinante, so genügen die Elementarteilerexponenten der Determinante der Form $A + \varrho B$, falls A eine reelle quadratische Form von verschwindender oder nichtverschwindender Determinante mit der Zahl q' als Wert der Charakteristik bedeutet und ϱ ein reeller variabler Parameter ist, der Ungleichheit:

$$q' \geqq s + \sum E\left(\frac{h}{2}\right) + \sum E\left(\frac{h'-1}{2}\right).$$

Hierbei ist $2s$ die Summe der Exponenten aller Elementarteiler, deren Basen für einen imaginären Wert des ϱ verschwinden; h durchläuft in der obigen Summe die Exponenten aller Elementarteiler, deren Basen von einem reellen, von Null verschiedenen Wert des ϱ annulliert werden, und h' nimmt die Werte der Exponenten aller Elementarteiler, die für $\varrho = 0$ verschwinden, an. $E\left(\frac{h}{2}\right)$ bedeutet die größte in $\frac{h}{2}$, $E\left(\frac{h'-1}{2}\right)$ die größte in $\frac{h'-1}{2}$ enthaltene ganze Zahl. Die Anzahl der für $\varrho = 0$ verschwindenden Elementarteiler ist gleich $n - l$, wenn B vom Range n, A vom Range l ist (A. Loewy, *Journ. f. Math.* **122**, 53 (1900)).

Hat A den Rang n, d. h. $|A| \neq 0$, so fällt in der Ungleichheit rechter Hand das Glied

$$\sum E\left(\frac{h'-1}{2}\right)$$

fort (vgl. F. Klein, *Bonner Diss.* (1868), *Math. Ann.* **23**, 561 (1884), A. Loewy, ebenda **52**, 588 (1899), *Gött. Nachr.* (1900), 298).

Für $q' = 0$ und $|A| \neq 0$ (*definite Form*) ergibt die Ungleichheit $s = 0$, $h = 1$ (*Satz von Weierstraß*); für $q' = 0$ und $|A| = 0$ (*semidefinite Form*) erhält man $s = 0$, $h = 1$, $h' = 1$ oder 2 (Gundelfinger, Suppl. zu Hesses *analyt. Geometrie d. Raumes*, S. 515, Gundelfinger-Dingeldey, *Vorles. a. d. analyt. Geom. d. Kegelschnitte*, Leipzig 1895, S. 67, vgl. auch den bei Muth, *Theorie der Elementarteiler*, S. 184 behandelten Fall, daß die Determinante $|A + \varrho B|$ identisch verschwindet

und A eine semidefinite Form ist). Wir formulieren noch den *Satz von Weierstraß* (*Ges. Werke* **1**, 233, **2**, 42, **3**, 139), der die im § 3 besprochenen Resultate über die Säkulargleichung als Spezialfall umfaßt:

Enthält eine Schar reeller quadratischer Formen eine definite Form, so besitzt die Determinante der Schar ausschließlich reelle Elementarteiler mit den Exponenten 1.

Mit der *reellen Transformation zweier Scharen reeller quadratischer Formen incinander* beschäftigt sich Muth, *Journ. f. Math.* **128**, 302 (1904).

Außer der bereits erwähnten Literatur sei noch die Monographie von Bromwich, *Quadratic forms and their classification by means of invariant factors*, Cambridge tracts (1906) genannt.

Die voraufgehenden Untersuchungen sind auch auf *Hermitesche Formen* ausdehnbar. Die Hermitesche Form

$$\sum_{i=1}^{i=n}\sum_{k=1}^{k=n} a_{ik}x_i\bar{x}_k,$$

bei der a_{ik} und a_{ki} konjugiert imaginär, a_{ii} reelle Größen sind, nimmt für konjugiert imaginäre Werte x_i und $\bar{x}_i$ nur reelle Werte an. Sie kann auf unendlich viele Arten durch zwei Substitutionen:

$$x_i = p_{i1}x_1' + p_{i2}x_2' + \cdots + p_{in}x_n'$$

und

$$\bar{x}_i = \bar{p}_{i1}\bar{x}_1' + \bar{p}_{i2}\bar{x}_2' + \cdots + \bar{p}_{in}\bar{x}_n' \qquad (i = 1, 2, \ldots, n),$$

wobei p_{ik} und $\bar{p}_{ik}$ konjugiert imaginäre Größen sind und die Determinante $|p_{ik}| \neq 0$ ist, in die *Normalform:*

$$x_1'\bar{x}_1' + x_2'\bar{x}_2' + \cdots + x_h'\bar{x}_h' - x_{h+1}'\bar{x}_{h+1}' - x_{h+2}'\bar{x}_{h+2}' - \cdots - x_l'\bar{x}_l'$$

transformiert werden.

Ebenso wie für reelle quadratische Formen gilt auch für Hermitesche Formen das *Trägheitsgesetz: Wie auch immer eine Hermitesche Form durch konjugiert imaginäre Substitutionen in die Normalform übergeführt wird, stets haben die zwei Zahlen h (Trägheitsindex) und l (Rang) dieselben unveränderlichen Werte* (Hermite, *Journ. f. Math.* **52**, 40 (1856), *Œuvres* **1**, 397, Frobenius, *Journ. f. Math.* **95**, 265 (1883)).

Die auf S. 127 besprochene Ungleichung

$$q' \geqq s + \sum E\left(\frac{h}{2}\right) + \sum E\left(\frac{h'-1}{2}\right)$$

bleibt auch unverändert gültig, wenn A und B anstatt reeller quadratischer Formen Hermitesche Formen sind (A. Loewy, *Journ. f. Math.* **122**, 67 (1900)). Wegen dieser Ungleichung vgl. auch Loewy, *Journ. f. Math.* **123**, 258 (1901), Bromwich, *Proc. Lond. M. S.* **32**, 349 (1900).

Auch die *alternierenden* bilinearen Formen sind eingehend untersucht worden. Die Determinante einer alternierenden bilinearen Form ist stets von geradem Rang (vgl. S. 64). Jede alternierende bilineare Form, deren Determinante den Rang $2r$ hat, kann durch kogrediente Substitutionen der zwei Variablenreihen in die Normalform

$$x_1'y_2' - x_2'y_1' + x_3'y_4' - x_4'y_3' + \cdots + x_{2r-1}'y_{2r}' - x_{2r}'y_{2r-1}'$$

übergeführt werden.

Der $2\lambda^{\text{te}}$ Elementarteiler einer alternierenden bilinearen Form

$$\sum_{i=1}^{i=n}\sum_{k=1}^{k=n} d_{ik}x_iy_k \qquad (d_{ik} = -d_{ki},\ d_{ii} = 0),$$

deren Koeffizienten d_{ik} ganze Zahlen oder ganze Funktionen eines Parameters ϱ sind, ist gleich dem $(2\lambda - 1)^{\text{ten}}$. Der $2\lambda^{\text{te}}$ Determinantenteiler $T_{2\lambda}$ (vgl. S. 105, Zeile 5) ist daher ein Quadrat (Verallgemeinerung des Cayleyschen Satzes auf S. 64).

Sind zwei alternierende bilineare Formen, deren Koeffizienten ganzzahlig oder ganze Funktionen eines Parameters ϱ sind, im Sinne der Festsetzungen auf S. 109 u. 110 äquivalent, so kann man die Formen stets kogredient ineinander transformieren.

Sind bei zwei Scharen bilinearer Formen $\varrho_1 A + \varrho_2 B$ und $\varrho_1 F + \varrho_2 G$ die Weierstraß-Kroneckerschen Bedingungen der Äquivalenz erfüllt und sind A und F symmetrisch, B und G alternierend, so sind die Scharen ebenso, wie wenn sie ausschließlich symmetrische oder ausschließlich alternierende Formen enthalten, kogredient ineinander transformierbar.

Die über alternierende Formen angegebenen Sätze stammen von Kronecker, *Monatsb. d. Berl. Akad.* (1874) 397, *Ges. Werke* **1**, 423 und Frobenius, *Journ. f. Math.* **86**, 165 u. 202, *Sitzungsb.*

d. Berl. Akad. (1896), 7. Vgl. ferner: E. v. Weber, *Sitzungsb. d. Bayer. Akad.* (1898), 369, Muth, *Theorie der Elementarteiler*, S. 134, *Journ. f. Math.* **122**, 89 (1900), Bromwich, *Am. J. math.* **23**, 235 (1901).

§ 10. Orthogonale Substitutionen und automorphe Transformationen quadratischer, Hermitescher und bilinearer Formen.

Bedeuten p_{ik} $(i,k=1,2,\ldots,n)$ n^2 Größen, zwischen denen die $\frac{n(n+1)}{2}$ Relationen:

$$p_{i1}^2 + p_{i2}^2 + \cdots + p_{in}^2 = 1 \qquad (i=1,2,\ldots,n),$$

$$p_{i1}p_{j1} + p_{i2}p_{j2} + \cdots + p_{in}p_{jn} = 0 \qquad (i \gtrless j;\ i,j=1,2,\ldots,n)$$

bestehen, so heißt die lineare homogene Substitution:

$$x_i = p_{i1}x_1' + p_{i2}x_2' + \cdots + p_{in}x_n' \qquad (i=1,2,\ldots,n)$$

orthogonal. Der Name stammt daher, weil die obigen Formeln für $n=2$, $n=3$ ein rechtwinkliges Koordinatensystem in ein ebensolches transformieren. Die Betrachtung orthogonaler Substitutionen ist von Euler (*Nov. Comm. Petrop.* **15**, 75 (1771) u. **20**, 189 u. 208 (1776), vgl. auch Jacobi, *Bemerkungen zu einer Abh. Eulers über die orthogonale Substitution* (Nachlaß), *Ges. Werke* **3**, 601) begonnen worden.

Die obigen $\frac{n(n+1)}{2}$ Gleichungen können gleichwertig durch die $\frac{n(n+1)}{2}$ Relationen (Euler, *Nov. Comm. Petrop.* **15**, 93)

$$p_{1i}^2 + p_{2i}^2 + \cdots + p_{ni}^2 = 1,$$

$$p_{1i}p_{1j} + p_{2i}p_{2j} + \cdots + p_{ni}p_{nj} = 0 \ (i \gtrless j)$$

ersetzt werden.

Ist P die Matrix der Koeffizienten, so wird eine orthogonale Substitution durch *jede der symbolischen Gleichungen* $P'P = E$ oder $PP' = E$ oder $P' = P^{-1}$ definiert; E ist die Einheitsmatrix und P' die zu P transponierte Matrix.

Die symbolische Gleichung $P'P = E$ besagt:

Jede orthogonale Substitution führt die spezielle quadratische Form:

$$x_1^2 + x_2^2 + \cdots + x_n^2$$

in sich selbst über.

Aus $P^{-1} = P'$ folgt, daß die *zu P reziproke Substitution:*

$$x_i' = p_{1i}x_1 + p_{2i}x_2 + \cdots + p_{ni}x_n \qquad (i = 1, 2, \ldots, n)$$

lautet (Lagrange, *Mécanique analytique, Œuvres* **12**, 205, vgl. Jacobi, *Ges. Werke* **3**, 603, ferner Cauchy, *Exercices de math.* **4** (1829), *Œuvres* (2) **9**, 193 u. 194).

Daher ist die *algebraische Adjungierte eines Elementes einer orthogonalen Substitution gleich dem Element selbst, wenn man es mit der Substitutionsdeterminante multipliziert* (Lagrange, *Œuvres* **12**, 206, Jacobi, *Journ. f. Math.* **12**, 9 (1834), *Ges. Werke* **3**, 201).

Die Determinante einer orthogonalen Substitution hat den Wert $+1$ *oder* -1 (Lagrange u. Jacobi, a. a. O.).

Eine orthogonale Substitution der Determinante $+1$ heißt *eigentlich orthogonal,* eine solche der Determinante -1 *uneigentlich orthogonal.*

Das Produkt zweier orthogonaler Substitutionen gleicher (ungleicher) Art ist eine eigentliche (uneigentliche) orthogonale Substitution.

Die charakteristische Gleichung

$$f(\varrho) = |\varrho\delta_{ik} - p_{ik}| = 0$$

einer orthogonalen Substitution mit den Koeffizienten p_{ik} *ist eine reziproke Gleichung* (Brioschi, *Journ. de math.* **19**, 253 (1854), *Faà di Bruno,* ebenda 304). Weiteren Aufschluß über die Natur von $f(\varrho)$ gibt der Satz von Frobenius (*Journ. f. Math.* **84**, 48 (1878)): *Die Elementarteiler der charakteristischen Funktion* $f(\varrho)$ *irgendeiner orthogonalen Substitution sind paarweise von gleichem Grade und für reziproke Werte Null, mit Ausnahme derjenigen, die für den Wert* $+1$ *und* -1 *verschwinden und einen ungeraden Exponenten haben. Ist umgekehrt* $f(\varrho)$ *ein Produkt von Elementarteilern der angegebenen Beschaffenheit, so gibt es eine orthogonale Substitution, deren charakteristische Funktion in die nämlichen Elementarteiler wie* $f(\varrho)$ *zerfällt.*

Aus dem allgemeinen Frobeniusschen Satz folgt das von einigen Autoren sogenannte *Stieltjessche Theorem* (*Acta math.* **6**, 319 (1885)) als Spezialfall (zu vgl. Voss, *Abh. d. Bayer. Akad.* (1890), 261): *Sind* a_{ik} *und* b_{ik} *die Elemente zweier orthogonaler Substitutionen gleichen Grades n, die beide*

zugleich eigentlich (uneigentlich) sind, und ist die Determinante mit dem allgemeinen Element $a_{ik} + b_{ik}$ Null, so sind auch alle ihre Unterdeterminanten vom $n - 1^{ten}$ Grade Null.

Ist die Determinante einer orthogonalen Substitution $|p_{ik}| = \varepsilon$ und ist -1 eine $q$fache und $+1$ eine $p$fache Wurzel von $f(\varrho) = 0$, so ist, wie sich aus dem Brioschischen Satz ergibt:

$$(-1)^q = \varepsilon, \quad (-1)^p = \varepsilon\,(-1)^n.$$

Eine unmittelbare Folge des Satzes von Frobenius ist das in der *Encycl. des sc. math.* **1**, 121 besonders hervorgehobene Resultat: Die ersten nichtverschwindenden Unterdeterminanten der Matrix

$$f(-1) = (-1)^n \,\|\, \delta_{ik} + p_{ik} \,\|\qquad (i, k = 1, 2, \ldots, n)$$

sind bei einer eigentlichen orthogonalen Substitution vom Grad $n - 2\lambda$, bei einer uneigentlichen vom Grad $n - 2\lambda - 1$, wobei λ eine ganze positive Zahl oder Null ist.

Ist T irgendeine alternierende Matrix, d. h. eine solche mit den Elementen $t_{ik} = -t_{ki}$ $(t_{ii} = 0)$ und verschwindet die Determinante $|\delta_{ik} + t_{ik}|$ $(\delta_{ii} = 1,\ \delta_{ik} = 0)$ nicht, so ist die *zu der Matrix:*

$$(E + T)^{-1} \cdot (E - T) = -E + 2\,(E + T)^{-1}$$

zugehörige Substitution eigentlich orthogonal. Ihre Koeffizienten lauten:

$$p_{ik} = \frac{2\beta_{ki}}{D}, \quad p_{ii} = \frac{2\beta_{ii} - D}{D};$$

hierbei ist

$$\beta_{ik} = \frac{\partial D}{\partial d_{ik}}$$

und D die Determinante $|d_{ik}| = |\delta_{ik} + t_{ik}|$. Die Determinante $|p_{ik} + \delta_{ik}|$ verschwindet nicht (*Cayleysche Formeln*, Cayley, *Journ. f. Math.* **32**, 120 (1846), *Coll. math. papers* **1**, 333, Frobenius, *Journ. f. Math.* **84**, 37 u. 50 (1878), sowie die Darstellung bei Baltzer, *Determinanten*, 5. Aufl., S. 190).

Umgekehrt: *Jede eigentliche orthogonale Substitution mit Koeffizienten p_{ik} und nichtverschwindender Determinante $|p_{ik} + \delta_{ik}|$ läßt sich in die obige Gestalt bringen, so daß die Koeffizienten p_{ik} als rationale Funktionen von $\frac{n(n-1)}{2}$ Parametern t_{ik} erscheinen.*

Die Cayleyschen Formeln versagen für alle uneigentlichen orthogonalen Substitutionen und für alle eigentlichen orthogonalen

Substitutionen, deren charakteristische Funktion -1 *zur Wurzel hat.* Jede orthogonale Substitution, die sich der Cayleyschen Darstellung entzieht, läßt sich in die Form eines Produktes einer durch die Cayleyschen Formeln gegebenen orthogonalen Substitution und einer Substitution der Form $x_i = \varepsilon_i x_i'$ $(i = 1, 2, \ldots, n)$ bringen, wobei die n Größen ε_i die positive oder negative Einheit bedeuten. Alle eigentlichen orthogonalen Substitutionen lassen sich auch aus den Cayleyschen Formeln durch Grenzübergang ableiten. Eine Behandlung der Ausnahmefälle bei Frobenius, *Journ. f. Math.* **84**, 42, Lipschitz, *Über Summen von Quadraten,* Bonn 1886, Kronecker, *Sitzungsb. d. Berl. Akad.* (1890), *Ges. Werke* **3**, 369, Prym, *Abh. d. Gött. Akad.* **38**, 1 (1892); vgl. auch die älteren Untersuchungen von Voss, *Math. Ann.* **13**, 320 (1878).

Sind sämtliche Elemente einer orthogonalen Substitution reell, so ist die Gleichung:

$$f(\varrho) = \begin{vmatrix} \varrho - p_{11} & -p_{12} & \cdots & -p_{1n} \\ -p_{21} & \varrho - p_{22} & \cdots & -p_{2n} \\ \vdots & & & \\ -p_{n1} & -p_{n2} & \cdots & \varrho - p_{nn} \end{vmatrix} = 0$$

nicht nur eine reziproke Gleichung, sondern sie hat die zwei weiteren Eigenschaften, nur Wurzeln vom absoluten Betrage 1 *zu besitzen und in lauter Elementarteiler mit den Exponenten* 1 *zu zerfallen, d. h. für jede m fache Wurzel verschwinden auch alle Unterdeterminanten* $n - 1^{ter}$, $n - 2^{ter}$, *bis* $n - m + 1^{ter}$ *Ordnung der linksstehenden Determinante* (Stickelberger, *Über reelle orthogonale Substitutionen,* Programm d. Züricher Polytechnikums, 1877, vgl. auch Frobenius, *Journ. f. Math.* **84**, 52, sowie die Darstellung bei Muth, *Theorie der Elementarteiler,* S. 173).

Die für reelle orthogonale Substitutionen angegebenen Resultate bleiben auch noch unverändert weiter bestehen, wenn man verlangt, die Substitution:

$$x_i = p_{i1} x_1' + p_{i2} x_2' + \cdots + p_{in} x_n' \qquad (i = 1, 2, \ldots, n)$$

soll mit ihrer konjugiert imaginären:

$$\bar{x}_i = \bar{p}_{i1} \bar{x}_1' + \bar{p}_{i2} \bar{x}_2' + \cdots + \bar{p}_{in} \bar{x}_n'$$

die Form

$$x_1\bar{x}_1 + x_2\bar{x}_2 + \cdots + x_n\bar{x}_n$$

mit konjugiert imaginären Variablen x_i und $\bar{x}_i$ in sich transformieren (Frobenius, *Journ. f. Math.* **95**, 265 (1883), A. Loewy, *C. R.* **123** (1896), *Nova Acta Leop.* **71**, 387 (1898), *Math. Ann.* **50**, 560 (1898)).

Sei T eine Matrix mit den Elementen t_{ik}, welche die Bedingung erfüllt, daß erstens die Determinante $|t_{ik} + \delta_{ik}| \neq 0$ ist und zweitens die mit der imaginären Einheit multiplizierte Matrix T eine Hermitesche Form definiert, d. h.

$$t_{ik} = c_{ik} + \sqrt{-1}\, d_{ik},\ c_{ik} = -c_{ki},\ d_{ik} = d_{ki} \text{ für } i \gtrless k,\ t_{ii} = \sqrt{-1}\, d_{ii},$$

wobei c_{ik}, d_{ik}, d_{ii} beliebige reelle Größen sind, so führt die zu der Matrix:

$$e^{\varphi\sqrt{-1}}(E+T)^{-1}(E-T) = e^{\varphi\sqrt{-1}}(-E + 2(E+T)^{-1})^{1})$$

zugehörige lineare homogene Substitution in Verbindung mit ihrer konjugiert imaginären die Form

$$x_1\bar{x}_1 + x_2\bar{x}_2 + \cdots + x_n\bar{x}_n$$

in sich über. Die Substitutionskoeffizienten lauten:

$$p_{ik} = \frac{2\beta_{ki}e^{\varphi\sqrt{-1}}}{D}, \quad p_{ii} = \frac{2\beta_{ii} - D}{D}\cdot e^{\varphi\sqrt{-1}},$$

$$\beta_{ik} = \frac{\partial D}{\partial d_{ik}}, \quad D = |d_{ik}| = |t_{ik} + \delta_{ik}|;$$

φ ist eine beliebige reelle Größe, δ_{ik} das Kroneckersche Symbol. *Diese Formeln — ein Analogon zu den Cayleyschen — stellen ohne Ausnahme* jede *Substitution dar, die mit ihrer konjugiert imaginären die Form*

$$x_1\bar{x}_1 + x_2\bar{x}_2 + \cdots + x_n\bar{x}_n$$

in sich transformiert (A. Loewy, *C. R.* **123** (1896), *Nova Acta Leop.* **71**, 393).

Die linearen homogenen Substitutionen P, die gemeinsam mit ihren konjugiert imaginären die Form $x_1\bar{x}_1 + x_2\bar{x}_2 + \cdots + x_n\bar{x}_n$ in sich transformieren, sind durch jede der Relationen $\bar{P}'P = E$ oder $P\bar{P}' = E$ oder $\bar{P}' = P^{-1}$ definiert. Die symbolische Glei-

1) e ist die Exponentielle.

chung $P\bar{P}' = E$ ist mit den $\frac{n(n+1)}{2}$ gewöhnlichen Gleichungen:

$$p_{i1}\bar{p}_{i1} + p_{i2}\bar{p}_{i2} + \cdots + p_{in}\bar{p}_{in} = 1 \qquad (i = 1, 2, \ldots, n),$$
$$p_{i1}\bar{p}_{j1} + p_{i2}\bar{p}_{j2} + \cdots + p_{in}\bar{p}_{jn} = 0 \qquad \begin{pmatrix} i \gtrless j, & i = 1, 2, \ldots, n \\ & j = 1, 2, \ldots, n \end{pmatrix}$$

gleichbedeutend. Mit diesen Formeln hängt die auch für die Theorie der Integralgleichungen (Fredholm, *Acta math.* **27**, 367 (1903)) wichtige Frage nach dem *Maximalwert des absoluten Betrages irgendeiner Determinante* innig zusammen. Bezeichnet man den Ausdruck $p_{i1}\bar{p}_{j1} + p_{i2}\bar{p}_{j2} + \cdots + p_{in}\bar{p}_{jn}$ mit P_{ij} $(i \gtrless j, i, j = 1, 2, \ldots, n)$, so ist der absolute Betrag *irgendeiner* Determinante $|P|$ gleich oder kleiner als die Quadratwurzel aus dem Produkt der n positiven Größen $P_{11}, P_{22}, \ldots, P_{nn}$, also $|P| \cdot |\bar{P}| \leqq P_{11} \cdot P_{22} \cdots P_{nn}$. Sind die Größen $P_{11} = P_{22} = \cdots = P_{nn} = 1$, so erreicht der absolute Betrag der Determinante $|P|$ dann und nur dann seinen Maximalwert 1, falls die $\frac{n(n-1)}{2}$ Größen P_{ij} $(i \gtrless j)$ verschwinden, d. h. P gemeinsam mit der konjugiert imaginären Substitution die Form $x_1\bar{x}_1 + x_2\bar{x}_2 + \cdots + x_n\bar{x}_n$ in sich transformiert (Hadamard, *Bull. sc. math.* (2) **17**, 240 (1893), Wirtinger, ebenda (2) **31**, 175 (1907), Davis, *Bull. Am. M. S.* **14**, 17 (1907), E. Fischer, Arch. f. Math. (3) **13**, 32 (1908), Pascal, *Determinanten*, S. 180).

Als Verallgemeinerung der orthogonalen Substitutionen hat zuerst Hermite (*Journ. f. Math.* **47**, 309 (1854), *Œuvres* **1**, 195) für $n = 3$ diejenigen linearen homogenen Transformationen

$$x_i = p_{i1}x_1' + p_{i2}x_2' + \cdots + p_{in}x_n' \qquad (i = 1, 2, \ldots, n)$$

betrachtet, die eine gegebene quadratische Form

$$\sum_{i=1}^{i=n} \sum_{k=1}^{k=n} a_{ik}x_i x_k$$

in sich, oder wie man mit Cayley (*Phil. Trans.* **148** (1858), *Coll. math. papers* **2**, 497) sagt, *automorph transformieren*.

Ist die Determinante der quadratischen Form ungleich Null, so hat die Determinante

$$|p_{ik}| \qquad (i, k = 1, 2, \ldots, n)$$

den Wert ± 1, ferner existiert, wie bereits Cayley (*Journ. f.*

Math. **50**, 288 (1855), *Coll. math. papers* **2**, 192; vgl. auch ebenda **2**, 497) zeigte, eine seinen für orthogonale Substitutionen gefundenen Formeln ähnliche Darstellung, schließlich bleibt der Satz von Frobenius über die Elementarteiler der charakteristischen Funktion $f(\varrho)$ gültig (Frobenius, *Journ. f. Math.* **84**, 41).

Als neueste Literatur über die Transformation einer quadratischen Form in sich sei angeführt: Percey F. Smith, *Trans. Am. M. S.* **6**, 1 (1905). Von Lehrbüchern sei verwiesen auf Muth, *Theorie der Elementarteiler*, S. 160, Bachmann, *Arithmetik der quadratischen Formen*, S. 395; schließlich vgl. man Franz Meyer, *Invariantentheorie in der Enzyklopädie der math. Wiss.* **1**, 333.

Soll eine *reelle quadratische Form reell* in sich transformiert werden, so gilt der Satz (A. Loewy, *Nova Acta Leop.* **71**, 396, 427, *Math. Ann.* **50**, 562, 569, *Gött. Nachr.* (1900), Bromwich, *Proc. Lond. M. S.* **32**, 350 (1900)):

Für jede lineare homogene Substitution P mit reellen Koeffizienten p_{ik}, die eine reelle quadratische Form von nichtverschwindender Determinante mit der Charakteristik q' (vgl. S. 123) in sich überführt, gilt die Ungleichheit:

$$q' \geqq s + \sum E\left(\frac{h}{2}\right).$$

$2s$ ist die Summe der Exponenten aller derjenigen Elementarteiler der charakteristischen Funktion $|\varrho\delta_{ik} - p_{ik}|$ der linearen homogenen Substitution P, welche für Größen, die nicht vom absoluten Betrage 1 sind, verschwinden. h durchläuft die Exponenten sämtlicher Elementarteiler der charakteristischen Funktion der Substitution P, welche für Größen vom absoluten Betrage 1 verschwinden. $E\left(\frac{h}{2}\right)$ bedeutet die größte in $\frac{h}{2}$ enthaltene ganze Zahl.

Ist $q' = 0$, d. h. die reelle quadratische Form ist definit, so wird $s = 0$, $h = 1$; in diesem Spezialfall ist *der Stickelbergersche Satz für reelle orthogonale Substitutionen* (vgl. S. 133) enthalten.

Die *automorphe Transformation einer bilinearen Form*

$$B = \sum_{i=1}^{i=n} \sum_{k=1}^{k=n} b_{ik} x_i y_k$$

mit kogredienten Variablen x_i, y_i $(i = 1, 2, \ldots, n)$ durch zwei kogrediente Substitutionen:

$$x_i = p_{i1} x_1' + p_{i2} x_2' + \cdots + p_{in} x_n' \qquad (i = 1, 2, \ldots, n),$$
$$y_i = p_{i1} y_1' + p_{i2} y_2' + \cdots + p_{in} y_n' \qquad (i = 1, 2, \ldots, n),$$

ist eingehend von A. Voss, *Abh. d. Bayer. Akad.* (1890), behandelt worden. Es handelt sich um die Gleichung $P'BP = B$. Als Lösungen führt bereits Cayley (*Coll. math. papers* **2**, 503, Abs. 14) an: $P = B^{-1}(B - T)(B + T')^{-1}B$, wobei T der Gleichung $B^{-1}T + (B')^{-1}T' = 0$ genügt (vgl. Voss, a. a. O., S. 305).

Die automorphe Transformation einer beliebigen bilinearen Form

$$B = \sum_{i=1}^{i=n} \sum_{k=1}^{k=n} b_{ik} \bar{x}_i x_k$$

mit konjugiert imaginären Variablen durch zwei Substitutionen:

$$x_i = p_{i1} x_1' + p_{i2} x_2' + \cdots + p_{in} x_n' \qquad (i = 1, 2, \ldots, n),$$
$$\bar{x}_i = \bar{p}_{i1} \bar{x}_1' + \bar{p}_{i2} \bar{x}_2' + \cdots + \bar{p}_{in} \bar{x}_n' \qquad (i = 1, 2, \ldots, n),$$

wobei stets g und $\bar{g}$ konjugiert imaginäre Größen bedeuten, behandelt A. Loewy, vgl. Zitat S. 136. Es handelt sich um die symbolische Gleichung $\bar{P}'BP = B$. Sind im besonderen b_{ik} und b_{ki} konjugiert imaginäre Größen, d h. ist

$$\sum_{i=1}^{i=n} \sum_{k=1}^{k=n} b_{ik} \bar{x}_i x_k$$

eine Hermitesche Form, so gilt für den Zusammenhang zwischen den Elementarteilern der Substitution P und der Charakteristik q' der Hermiteschen Form B die nämliche Ungleichheit wie bei der reellen automorphen Transformation einer reellen quadratischen Form, ferner finden ähnliche ausnahmslos gültige Formeln wie bei der automorphen Transformation der Hermiteschen Einheitsform $x_1 \bar{x}_1 + x_2 \bar{x}_2 + \cdots + x_n \bar{x}_n$ statt. Geometrische Anwendungen der automorphen Transformation Hermitescher Formen bei Study, *Math. Ann.* **60**, 321 (1905).

§ 11. Abgeleitete Matrices. Frankesche und Sylvestersche Sätze. Laplacescher und Jacobischer Satz. Potenztransformation. Produkttransformation. Lineare homogene Substitutionen, die eine Determinante in sich überführen.

Zu jeder quadratischen Matrix $A = \| a_{ik} \|$ $(i, k = 1, 2, \ldots, n)$ gehören $\binom{n}{m}^2$ Determinanten

$$\sum \pm a_{g_1 h_1} a_{g_2 h_2} \ldots a_{g_m h_m} \quad (g_1 < g_2 < \cdots < g_m;\ h_1 < h_2 < \cdots < h_m)$$

(vgl. S. 56). Um diese Unterdeterminanten der Determinante $|A| = |a_{ik}|$ $(i, k = 1, 2, \ldots, n)$ passend bezeichnen zu können, denken wir uns die $\lambda = \binom{n}{m}$ Kombinationen der Zahlen $1, 2, \ldots, n$ zu m in eine beliebige, aber fest gewählte Reihenfolge gebracht. Ist in der eingeführten Anordnung $g_1, g_2, \ldots, g_m$ die G^{te} und $h_1, h_2, \ldots, h_m$ die H^{te} Kombination, so sei die Determinante

$$\sum \pm a_{g_1 h_1} a_{g_2 h_2} \ldots a_{g_m h_m}$$

mit $C^{(m)}_{GH}$ bezeichnet. Aus den $\lambda^2 = \binom{n}{m}^2$ Determinanten $C^{(m)}_{GH}$ sei die Matrix

$$\begin{Vmatrix} C^{(m)}_{11} & C^{(m)}_{12} & \ldots & C^{(m)}_{1\lambda} \\ C^{(m)}_{21} & C^{(m)}_{22} & \ldots & C^{(m)}_{2\lambda} \\ \vdots & \vdots & & \vdots \\ C^{(m)}_{\lambda 1} & C^{(m)}_{\lambda 2} & \ldots & C^{(m)}_{\lambda\lambda} \end{Vmatrix}$$

gebildet, die wir mit $C_m(A)$ bezeichnen.

$C_m(A)$ heißt *das m^{te} abgeleitete System der Matrix A* (Kronecker, *Vorl. üb. Determinanten,* S. 320). Bei natürlicher Anordnung der Zahlen $1, 2, \ldots, n$ ist das erste abgeleitete System $C_1(A)$ mit A identisch, das letzte abgeleitete System $C_n(A)$ ist die Determinante $|A|$. Nach Henry J. St. Smith (*Proc. Royal Soc.* **12** (1864), *Coll. math. papers* **1**, 412), der sich zuerst mit den abgeleiteten Systemen beschäftigt hat, heißt $C_m(A)$ die *m^{te} Concomitante von A;* später bezeichnete er (*Coll. math. papers* **2**, 625) $C_m(A)$ als die *$m-1^{te}$ adjungierte Matrix* von $C_1(A)$. Nach Bachmann, *Die Arithmetik*

der quadratischen Formen, Leipzig 1898, S. 389, würde man die Systeme $C_m(A)$ *begleitende Matrices* nennen. Dickson, *Linear groups, Teubners Samml.* **6**, Leipzig 1901, S. 146, verwendet englischem Sprachgebrauch gemäß (vgl. etwa Spottiswoode, *Journ. f. Math.* **51**, 350) die Bezeichnung *„the mth compound"*. C. Stephanos, *Journ. de math.* (5) **6**, 73 (1900) nennt $C_m(A)$ *le produit bialterné de m formes A.* In der Theorie der linearen homogenen Differentialgleichungen begegnet man $C_m(A)$ unter dem Namen *„$n - m^{tes}$ assoziiertes System"* (vgl. L. Schlesinger, *Handbuch der Theorie der linearen Differentialgleichungen,* Leipzig 1897, $\mathbf{2}_1$, 128, A. Loewy, *Sitzungsb. d. Bayer. Akad.* **32**, 3 (1902)). Für die abgeleiteten Systeme gelten folgende Lehrsätze:

I. Sind A_1 und A_2 zwei beliebige Matrices gleiche Grades n, so ist das Produkt von $C_m(A_1)$ und $C_m(A_2)$ gleich der m^{ten} abgeleiteten Matrix des Produktes $A_1 A_2$, also

$$C_m(A_1)\, C_m(A_2) = C_m(A_1 A_2).$$

II. Ist A die Einheitsmatrix, so ist auch $C_m(A)$ die Einheitsmatrix.

Aus Lehrsatz I und II folgt:

III. Die m^{te} abgeleitete Matrix der reziproken Matrix von A ist gleich der reziproken Matrix der m^{ten} Abgeleiteten von A, also

$$C_m(A^{-1}) = (C_m(A))^{-1}.$$

IV. Die m^{te} abgeleitete Matrix der transponierten Matrix von A ist gleich der transponierten Matrix der m^{ten} abgeleiteten Matrix von A, also

$$C_m(A') = (C_m(A))'.$$

V. Man erhält die $\binom{n}{m}$ Wurzeln der charakteristischen Gleichung von $C_m(A)$, d. h. der Gleichung $|\varrho E - C_m(A)| = 0$, indem man die n Wurzeln der charakteristischen Gleichung von A zu je m kombiniert und miteinander multipliziert (W. H. Metzler, *Am. J. math.* **16**, 145 (1894), G. Rados, *Math. Ann.* **48**, 417 (1897), W. Burnside, *Quart. J.* **32**, 84 (1901)).

Aus Lehrsatz V folgt der sogenannte Satz von Franke· (*Journ. f. Math.* **61**, 353 (1863)): Die Determinante von $C_m(A)$ ist die $\binom{n-1}{m-1}^{te}$ Potenz der Determinante von A, also

$$|C_m(A)| = |A|^{\binom{n-1}{m-1}}.$$

Dieser Satz ist aber bereits vor Franke von Spottiswoode (*Journ. f. Math.* **51**, 360 (1856)) ausgesprochen worden. Das von Spottiswoode a. a. O. gegebene Zitat (*Phil. Mag.* (1851)) bezieht sich auf eine Arbeit von Sylvester (*Coll. math. papers* **1**, 249), vgl. auch die Note von Baker, ebenda S. 650, ferner die Anmerkung von Borchardt zur Arbeit von Franke, *Journ. f. Math.* **61**, 355. Der fälschlich nach Franke genannte Satz ist nämlich nur ein Spezialfall eines *allgemeinen von Sylvester* a. a. O. *gegebenen Theorems,* das folgendermaßen lautet:

Aus der Matrix $C_{k+m}(A)$ sei eine Determinante gebildet, deren Elemente nur alle $\sigma^2 = \binom{n-k}{m}^2$ Determinanten

$$C_{GH}^{(k+m)} = \sum \pm a_{11}\, a_{22} \dots a_{kk}\, a_{g_1 h_1}\, a_{g_2 h_2} \dots a_{g_m h_m}$$

sind, bei denen $g_1, g_2, \dots, g_m$ und $h_1, h_2, \dots, h_m$ jede Kombination der Zahlen $k+1, k+2, \dots, n$ zu m bedeuten. Die fragliche Determinante $\left| C_{GH}^{(k+m)} \right|_{(G,H=1,2,\dots,\sigma)}$ hat den Wert

$$A_k^{\binom{n-k-1}{m}} \cdot |A|^{\binom{n-k-1}{m-1}},$$

wobei

$$A_k = \sum \pm a_{11}\, a_{22} \dots a_{kk}$$

und

$$|A| = \sum \pm a_{11}\, a_{22} \dots a_{nn} \quad (k+m \leqq n).$$

Für $k=0$, $A_0 = 1$ hat man den sogenannten Frankeschen Satz; $m=1$, $\binom{n-k-1}{0} = 1$ liefert den spezielleren Sylvesterschen Satz auf S. 62. Betreffs des obigen allgemeinen Sylvesterschen Satzes vgl. man: Picquet, *J. éc. polyt.*, Cah. **45**, 216 (1878), Sylvester, *Journ. f. Math.* **88**, 52 (1880), Netto, ebenda **114**, 351 (1895)[1]), Nanson, ebenda **122**, 183 (1900), E. Müller, *Ztschr. f. Math. u. Physik* **44**, 39 (1899), Pascal, *Determinanten,* S. 93, Scott, *The theory of determinants,* S. 67, 68.

Zu einer neuen Definition der Systeme $C_m(A)$ führt folgende Betrachtung: Wir ordnen der Matrix A die m kogredienten linearen homogenen Substitutionen

---

1) Aus der Formel auf S. 352 und der Ende S. 351 folgt die Formel des Textes.

$$\begin{aligned} y_1^{(i)} &= a_{11}x_1^{(i)} + a_{12}x_2^{(i)} + \cdots + a_{1n}x_n^{(i)}, \\ y_2^{(i)} &= a_{21}x_1^{(i)} + a_{22}x_2^{(i)} + \cdots + a_{2n}x_n^{(i)}, \\ &\vdots \\ y_n^{(i)} &= a_{n1}x_1^{(i)} + a_{n2}x_2^{(i)} + \cdots + a_{nn}x_n^{(i)} \end{aligned}$$
$$(i = 1, 2, \ldots, m)$$

zu. Versteht man unter $\mathfrak{X}$ die quadratische Matrix

$$\left\| \begin{matrix} x_1^{(1)} & x_1^{(2)} & \ldots & x_1^{(m)} & \overbrace{0} & \overbrace{0} & \ldots & 0 \\ x_2^{(1)} & x_2^{(2)} & \ldots & x_2^{(m)} & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ x_n^{(1)} & x_n^{(2)} & \ldots & x_n^{(m)} & 0 & 0 & \ldots & 0 \end{matrix} \right\|$$

($n-m$ Kolonnen)

und unter Y die analog gebaute, so lassen sich die obigen Gleichungen für die m kogredienten Substitutionen in die eine symbolische Gleichung

$$Y = AX \tag{1}$$

zusammenfassen (Verallgemeinerung des auf S. 82 angegebenen Resultates).

Aus der Gleichung (1) folgt auf Grund des Lehrsatzes I

$$C_m(Y) = C_m(A)\, C_m(X). \tag{2}$$

Sind $l_1, l_2, \ldots, l_m$ irgend m untereinander verschiedene der n Zahlen $1, 2, \ldots, n$, die in der Aufeinanderfolge $l_1 < l_2 < \cdots < l_m$ angeordnet sein sollen, so bilde man für jede mögliche Wahl der Zahlen $l_1, l_2, \ldots, l_m$ die Determinante

$$\begin{vmatrix} x_{l_1}^{(1)} & x_{l_1}^{(2)} & \ldots & x_{l_1}^{(m)} \\ x_{l_2}^{(1)} & x_{l_2}^{(2)} & \ldots & x_{l_2}^{(m)} \\ \vdots & \vdots & & \vdots \\ x_{l_m}^{(1)} & x_{l_m}^{(2)} & \ldots & x_{l_m}^{(m)} \end{vmatrix}$$

Ist in der zu Beginn des Paragraphen für die λ Kombinationen der Zahlen $1, 2, \ldots, n$ zu m eingeführten Reihenfolge $l_1, l_2, \ldots, l_m$

die L^{te}, so sei die zuletzt hingeschriebene Determinante mit X_L bezeichnet. Analog seien die λ Determinanten

$$Y_L = \begin{vmatrix} y_{i_1}^{(1)} & y_{i_1}^{(2)} & \dots & y_{i_1}^{(m)} \\ y_{i_2}^{(1)} & y_{i_2}^{(2)} & \dots & y_{i_2}^{(m)} \\ \vdots & \vdots & & \vdots \\ y_{i_m}^{(1)} & y_{i_m}^{(2)} & \dots & y_{i_m}^{(m)} \end{vmatrix} \qquad (L = 1, 2, \dots, \lambda)$$

definiert. Die symbolische Gleichung (2) kann dann in die λ gewöhnlichen Gleichungen

$$(3) \qquad Y_L = \sum_{J=1}^{J=\lambda} C_{LJ}^{(m)} X_J \qquad (L = 1, 2, \dots, \lambda)$$

umgesetzt werden.

Die Gleichung (3) besagt:

Transformiert man die m Variablensysteme

$$y_i^{(1)}, y_i^{(2)}, \dots, y_i^{(m)} \qquad (i = 1, 2, \dots, n)$$

kogredient durch m lineare homogene Substitutionen der nämlichen Matrix A, so erfahren die λ Determinanten Y_J $(J = 1, 2, \dots, \lambda)$ eine lineare homogene Substitution, die durch die Matrix $C_m(A)$ gegeben ist. A. Hurwitz, *Math. Ann.* **45**, 392 (1894) nennt daher die durch die Formeln (3) bestimmte lineare homogene Substitution die *m^{te} Determinantentransformation.* Die Determinanten X_J sind von A. Clebsch (*Math. Ann.* **5**, 427 (1872)) als selbständige Variablen eingeführt worden. Die fraglichen Determinanten lassen sich auch geometrisch deuten. Erklärt man

$$x_1^{(i)}, x_2^{(i)}, \dots, x_n^{(i)} \qquad (i = 1, 2, \dots, m)$$

als homogene Koordinaten von m Punkten, so stellen $X_1, X_2, \dots, X_\lambda$ die Koordinaten der durch die m Punkte bestimmten ebenen Mannigfaltigkeit dar.

Wird die bilineare Form

$$A = \sum_{i=1}^{i=n} \sum_{k=1}^{k=n} a_{ik} x_i y_k$$

durch die zwei linearen homogenen Substitutionen

$$P: \qquad x_i = p_{i1} x_1' + p_{i2} x_2' + \dots + p_{in} x_n' \qquad (i = 1, 2, \dots, n),$$

$$Q: \qquad y_i = q_{i1} y_1' + q_{i2} y_2' + \dots + q_{in} y_n' \qquad (i = 1, 2, \dots, n)$$

in die bilineare Form

$$B = \sum_{i=1}^{i=n} \sum_{k=1}^{k=n} b_{ik} x_i' y_k'$$

übergeführt, besteht also die symbolische Gleichung $P'AQ = B$, so folgt aus ihr:

$$(C_m(P))' C_m(A) C_m(Q) = C_m(B).$$

Diese Gleichung besagt: die m^{ten} Determinantentransformationen von P und Q führen die m^{ten} abgeleiteten bilinearen Formen von A und B ineinander über. Hieraus folgt im besonderen, daß, wenn P orthogonal ist, dies auch für $C_m(P)$ zutrifft.

Ist A eine quadratische bzw. Hermitesche Form, so ist auch die m^{te} abgeleitete Form $C_m(A)$ von der nämlichen Beschaffenheit. Mit den Abgeleiteten einer quadratischen Form beschäftigt sich Rados, *Verh. des ersten internationalen Math.-Kongr.* (1897), 163.

Wir ordnen der Unterdeterminante

$$C_{GH}^{(m)} = \sum \pm a_{g_1 h_1} a_{g_2 h_2} \dots a_{g_m h_m}$$

von $|A|$ ihre algebraische Adjungierte

$$\frac{\partial^m |A|}{\partial a_{g_1 h_1} \partial a_{g_2 h_2} \dots \partial a_{g_m h_m}}$$

zu und bezeichnen sie mit $\Gamma_{GH}^{(m)}$. Aus den λ^2 Größen $\Gamma_{GH}^{(m)}$ bilden wir die Matrix

$$\left\| \begin{matrix} \Gamma_{11}^{(m)} & \Gamma_{12}^{(m)} & \dots & \Gamma_{1\lambda}^{(m)} \\ \Gamma_{21}^{(m)} & \Gamma_{22}^{(m)} & \dots & \Gamma_{2\lambda}^{(m)} \\ \vdots & \vdots & & \vdots \\ \Gamma_{\lambda 1}^{(m)} & \Gamma_{\lambda 2}^{(m)} & \dots & \Gamma_{\lambda\lambda}^{(m)} \end{matrix} \right\|;$$

sie sei mit $\Gamma_m(A)$ bezeichnet.

Die Lehrsätze I—IV bleiben unverändert gültig, wenn man überall C_m durch Γ_m ersetzt. Dem Satz V stellt sich der Lehrsatz V′ zur Seite: Man erhält die $\binom{n}{n-m}$ Wurzeln der charakteristischen Gleichung von $\Gamma_m(A)$, d. h. der Gleichung

$$|\varrho E - \Gamma_m(A)| = 0,$$

indem man die n Wurzeln der charakteristischen Gleichung von A zu je $n - m$ kombiniert und miteinander multipliziert. Hieraus folgt

$$|\Gamma_m(A)| = |A|^{\binom{n-1}{m}}.$$

Mittels der Elemente der Matrices $C_m(A)$ und $\Gamma_m(A)$ kann man *den Laplaceschen Zerlegungssatz* und das sich ihm anschließende Theorem auf S. 58 durch die Gleichungen

$$\sum_{J=1}^{J=\lambda} C_{GJ}^{(m)} \Gamma_{HJ}^{(m)} = |A| \delta_{GH} \qquad (G, H = 1, 2, \ldots, \lambda)$$

ausdrücken; δ_{GH} ist hierbei das Kroneckersche Symbol, das für $G = H$ den Wert 1 hat, sonst verschwindet. Die angegebenen λ^2 Gleichungen lassen sich in eine einzige symbolische Gleichung, nämlich

$$C_m(A)\,\Gamma_m(A') = |A|, \tag{1}$$

zusammenfassen; $|A|$ ist als Diagonalmatrix aufzufassen, deren in der Diagonale stehende Elemente sämtlich den Wert $|A|$ haben.

Da nach Lehrsatz III

$$C_m(A)\,C_m(A^{-1}) = E \tag{2}$$

ist, folgt aus (1) und (2) die Kroneckersche Gleichung (Kronecker, *Monatsb. d. Berl. Akad.* 1882, *Ges. Werke* **2**, 392, *Vorl. über Determinanten*, S. 334):

$$\Gamma_m(A) = |A|\, C_m(A'^{-1}).$$

Bezeichnet man die nach gewöhnlichem Sprachgebrauch *adjungierte Matrix* $\|A_{ik}\|$ $(i, k = 1, 2, \ldots, n)$ mit $\mathrm{adj.}(A)$ (vgl. S. 60, 61), so ist (vgl. S. 87):

$$A'^{-1} = \frac{\mathrm{adj.}(A)}{|A|};$$

mithin wird:

$$C_m(A'^{-1}) = C_m\left(\frac{\mathrm{adj.}(A)}{|A|}\right) = \frac{1}{|A|^m}\, C_m(\mathrm{adj.}(A)).$$

Setzt man den für $C_m(A'^{-1})$ gefundenen Wert in die Kroneckersche Gleichung ein, so ergibt sich:

$$C_m(\mathrm{adj.}(A)) = |A|^{m-1} \Gamma_m(A).$$

Diese Gleichung ist nichts anderes als der Ausdruck der auf S. 61 gegebenen *Jacobischen Formel.* Sie gilt auch noch, falls die Determinante $|A|$ verschwindet; bei der Ableitung ist dies zunächst infolge der Verwendung von A^{-1} ausgeschlossen worden.

Wir ersetzen in der Kroneckerschen Gleichung die Matrix A durch $\Gamma_h(B)$, wobei B eine beliebige Matrix n^{ten} Grades ist. Verwendet man die Kroneckersche Gleichung noch ein zweites Mal, so erhält man:

$$\begin{aligned}\Gamma_m(\Gamma_h(B)) &= |\Gamma_h(B)|\, C_m(\Gamma_h(B'^{-1})) \\ &= |B|^{\binom{n-1}{h}} C_m(|B'^{-1}|\, C_h(B)) \\ &= \frac{|B|^{\binom{n-1}{h}}}{|B|^m} C_m(C_h(B)).\end{aligned}$$

Neben die Gleichung:

$$\Gamma_m(\Gamma_h(B)) = |B|^{\binom{n-1}{h}-m} C_m(C_h(B)) \tag{3}$$

stellt sich die Gleichung:

$$\Gamma_m(C_h(B)) = |B|^{\binom{n-1}{h-1}-m} C_m(\Gamma_h(B)). \tag{4}$$

Die Gleichung (4) folgt ähnlich wie (3) aus der Kroneckerschen Relation, wenn man in ihr für A die Matrix $C_h(B)$ einführt. Die Relation (3) wird ebenfalls (vgl. S. 139) als *Frankescher Satz* bezeichnet (Franke, *Journ. f. Math.* **61**, 355 (1863), Baltzer, *Determinanten,* S. 68 u. 69). Die Gleichungen (3) und (4) gelten auch noch, falls die Matrix B eine verschwindende Determinante hat; bei der Herleitung ist dies infolge der Benützung der Kroneckerschen Relation zunächst auszuschließen.

Sind A und B zwei beliebige Matrices gleichen Grades, so ergibt sich auf Grund der Gleichung (1):

$$C_h(A)\,\Gamma_h(B')\,C_h(B)\,\Gamma_h(A') = |A|\,|B|. \tag{5}$$

Führt man zur Abkürzung die Matrices $T = C_h(A)\,\Gamma_h(B')$ und $U = C_h(B)\,\Gamma_h(A')$ ein, so hat man die symbolische Gleichung:

$$TU = |A|\,|B|. \tag{6}$$

Für die Determinanten erhält man:

$$(7) \qquad |T| = |C_h(A)|\,|\Gamma_h(B')| = |A|^{\binom{n-1}{h-1}}\,|B|^{\binom{n-1}{h}},$$

$$(8) \qquad |U| = |C_h(B)|\,|\Gamma_h(A')| = |B|^{\binom{n-1}{h-1}}\,|A|^{\binom{n-1}{h}},$$

$$(9) \qquad |TU| = |A|^{\binom{n}{h}}\,|B|^{\binom{n}{h}}$$

(Sylvester, *Phil. Mag.* (1851), *Coll. math. papers* **1**, 253, vgl. ebenda die Note von Baker, S. 649, Baltzer, *Determinanten,* S. 36, Pascal, *Determinanten,* S. 111, wegen einer Verallgemeinerung vgl. E. Müller, *Ztschr. f. Math. u. Phys.* **44**, 31 (1899)).

Aus der Gleichung (6) folgt:

$$C_m(TU) = C_m(|A|\,|B|) \text{ und daher}$$
$$C_m(T) = |A|^m\,|B|^m\,C_m(U^{-1}).$$

Mit Hilfe der Kroneckerschen Relation ergibt sich weiter:

$$C_m(T) = |A|^m\,|B|^m\,|U^{-1}|\,\Gamma_m(U') \text{ und schließlich}$$

$$(10) \qquad C_m(T) = |A|^{m-\binom{n-1}{h}}\,|B|^{m-\binom{n-1}{h-1}}\,\Gamma_m(U').$$

Die Formel (10) wird als *Satz von Picquet* (*C. R.* **86**, 1119 (1878), *J. éc. polyt.,* Cah. **45**, 238 (1878), Pascal, *Determinanten,* S. 112) bezeichnet. Wählt man für die Matrix A bezw. B die Einheitsmatrix E, so ergibt Formel (10) die Formeln (4) bezw. (3).

Als Literatur über die abgeleiteten Systeme sei noch angeführt: Vahlen, *Enzykl. der math. Wiss.* **1**, 594, J. Schur, *Berl. Diss.* (1901), A. Loewy, *Trans. Am. M. S.* **5**, 64 (1904), sowie die ergänzenden Bemerkungen, ebenda **6**, 507 (1905).

Die Gleichung (5) geht, falls für B die Einheitsmatrix E gesetzt wird, in die Laplacesche Gleichung (1) über. Eine andersartige Erweiterung der Laplaceschen Entwicklung stammt von E. Netto, *Journ. f. Math.* **114**, 348 (1895). Vgl. Nanson, ebenda **122**, 182 (1900), E. Müller, *Ztschr. f. Math. u. Phys.* **44**, 33 (1899), Pascal, *Determinanten,* S. 97, Scott, *The theory of determinants,* S. 70.

Zu jeder Matrix A läßt sich auf folgende Weise eine neue Matrix bilden. Sei

$$x_i = a_{i1}\xi_1 + a_{i2}\xi_2 + \cdots + a_{in}\xi_n \qquad (i = 1, 2, \ldots, n)$$

eine zu A zugehörige lineare homogene Substitution. Man bilde sich sämtliche verschiedene Produkte:

$$x_{i_1} x_{i_2} \ldots x_{i_r} \qquad (i_1, i_2, \ldots, i_r = 1, 2, \ldots, n);$$

ihre Anzahl ist gleich $\binom{n+r-1}{r}$, nämlich gleich der Zahl, die angibt, auf wieviel Arten man n Zahlen zu r mit Wiederholung kombinieren kann. Diese $\binom{n+r-1}{r}$ Produkte transformieren sich linear und homogen in die $\binom{n+r-1}{r}$ Produkte $\xi_{i_1}\xi_{i_2}\ldots\xi_{i_r}$. Denkt man sich die Produkte $x_{i_1}x_{i_2}\ldots x_{i_r}$ willkürlich, aber, wenn gewählt, fest in eine bestimmte Reihenfolge gebracht und die Produkte $\xi_{i_1}\xi_{i_2}\ldots\xi_{i_r}$ in der nämlichen Reihenfolge angeordnet, so heißt die sich ergebende eindeutig bestimmte lineare homogene Substitution nach A. Hurwitz (*Math Ann.* **45**, 390 (1894)) die *r^{te} Potenztransformation von A* oder anknüpfend an Sylvester die *induzierte Substitution r^{ten} Grades* (Franklin, *Am. J. math.* **16**, 205 (1894), G. Rados, *Math. u. naturw. Ber. aus Ungarn* **16**, 241 (1898)). Die Matrix der r^{ten} Potenztransformation von A wird nach A. Hurwitz mit $P_r(A)$ bezeichnet.

Für die Operation $P_r(A)$ gelten folgende Sätze:

I. Sind A_1 und A_2 zwei beliebige Matrices gleichen Grades, so ist das Produkt von $P_r(A_1)$ und $P_r(A_2)$ gleich der r^{ten} Potenztransformation des Produktes von A_1 und A_2, also:

$$P_r(A_1)\,P_r(A_2) = P_r(A_1 A_2).$$

II. Ist A die Einheitsmatrix, so ist auch $P_r(A)$ die Einheitsmatrix.

III. Die r^{te} Potenztransformation der reziproken Matrix von A ist gleich der reziproken Matrix der r^{ten} Potenztransformation von A, also:

$$P_r(A^{-1}) = (P_r(A))^{-1}.$$

IV. Die r^{te} Potenztransformation der transponierten Matrix von A ist gleich der transponierten Matrix der r^{ten} Potenztransformation von A, also:

$$P_r(A') = (P_r(A))'.$$

V. Man erhält die $\binom{n+r-1}{r}$ Wurzeln der charakteristischen Gleichung von $P_r(A)$, d. h. der Gleichung $|\varrho E - P_r(A)| = 0$, indem man die sämtlichen $\binom{n+r-1}{r}$ Produkte der n Wurzeln der charakteristischen Gleichung von A zu je r bildet (Franklin, *Am. J. math.* **16**, 205 (1894), G. Rados, *Math. u. naturw. Ber. aus Ungarn* **16**, 244 (1898), W. Burnside, *Quart. J.* **32**, 83 (1901), J. Schur, *Berl. Diss.* (1901), 17).

Aus Lehrsatz V folgt: Die Determinante von $P_r(A)$ ist die $\binom{n+r-1}{r-1}^{\text{te}}$ Potenz der Determinante von A, also

$$|P_r(A)| = |A|^{\binom{n+r-1}{r-1}}.$$

(G. v. Escherich, *Monatsh. f. Math.* **3**, 80 (1892), A. Hurwitz, *Math. Ann.* **45**, 391 (1894), W. Anissimoff, ebenda **51**, 388 (1899); für $r = 2$ wird dieser Determinantensatz von Pascal, *Determinanten*, S. 104 als Scholtzscher Satz bezeichnet.) Weiteres über die Potenztransformation folgt unten bei der Produkttransformation.

Sind eine Reihe beliebiger quadratischer Matrices $A_1, A_2, \ldots, A_r$ der Grade $n_1, n_2, \ldots, n_r$ gegeben, so läßt sich aus ihnen eine neue Matrix vom Grade $n_1 n_2 \ldots n_r$ konstruieren, die mit $A_1 \times A_2 \times \cdots \times A_r$ bezeichnet wird. Sie wird am einfachsten auf folgende Weise definiert: Man führe die zu den r Matrices

$$A_\nu = \| a_{ik}^{(\nu)} \| \qquad (i, k = 1, 2, \ldots, n_\nu;\ \nu = 1, 2, \ldots, r)$$

zugehörigen linearen homogenen Substitutionen:

$$A_1: \quad x_{i_1}^{(1)} = \sum_{s_1=1}^{s_1=n_1} a_{i_1 s_1}^{(1)} \xi_{s_1}^{(1)} \qquad (i_1 = 1, 2, \ldots, n_1),$$

$$A_2: \quad x_{i_2}^{(2)} = \sum_{s_2=1}^{s_2=n_2} a_{i_2 s_2}^{(2)} \xi_{s_2}^{(2)} \qquad (i_2 = 1, 2, \ldots, n_2),$$

$$\vdots \qquad\qquad \vdots$$

$$A_r: \quad x_{i_r}^{(r)} = \sum_{s_r=1}^{s_r=n_r} a_{i_r s_r}^{(r)} \xi_{s_r}^{(r)} \qquad (i_r = 1, 2, \ldots, n_r)$$

ein. Durch Multiplikation erhält man:

$$(1) \qquad x_{i_1}^{(1)} x_{i_2}^{(2)} \dots x_{i_r}^{(r)} = \sum_{k_1 k_2 \dots k_r} a_{i_1 k_1}^{(1)} a_{i_2 k_2}^{(2)} \dots a_{i_r k_r}^{(r)} \xi_{k_1}^{(1)} \xi_{k_2}^{(2)} \dots \xi_{k_r}^{(r)};$$

hierbei ist dem Index i_1 jede der Zahlen $1, 2, \dots, n_1$, dem Index i_2 jede der Zahlen $1, 2, \dots, n_2$ usw., dem Index i_r jede der Zahlen $1, 2, \dots, n_r$ beizulegen. Ebenso ist in der Summe rechts für k_1 jede der Zahlen $1, 2, \dots, n_1$, für k_2 jede der Zahlen $1, 2, \dots, n_2$ usw., für k_r jede der Zahlen $1, 2, \dots, n_r$ zu setzen. Die $n_1 n_2 \dots n_r$ Produkte $x_{i_1}^{(1)} x_{i_2}^{(2)} \dots x_{i_r}^{(r)}$ sind offenbar lineare homogene Funktionen der $n_1 n_2 \dots n_r$ Produkte $\xi_{i_1}^{(1)} \xi_{i_2}^{(2)} \dots \xi_{i_r}^{(r)}$.

Hat man für die $n_1 n_2 \dots n_r$ Produkte $x_{i_1}^{(1)} x_{i_2}^{(2)} \dots x_{i_r}^{(r)}$ eine beliebige, aber fest gewählte Reihenfolge und für die $n_1 n_2 \dots n_r$ Produkte $\xi_{i_1}^{(1)} \xi_{i_2}^{(2)} \dots \xi_{i_r}^{(r)}$ die nämliche Anordnung eingeführt, so definiert das Gleichungssystem (1) eindeutig eine bestimmte lineare homogene Substitution in $n_1 n_2 \dots n_r$ Variablen; sie heißt nach A. Hurwitz (*Math. Ann.* **45**, 388 (1894)) die *Produkttransformation* von $A_1, A_2, \dots, A_r$ und wird nach ihm mit $A_1 \times A_2 \times \dots \times A_r$ bezeichnet. C. Stephanos (*Journ. de math.* (5) **6**, 73 (1900)), der die Produkttransformation sehr eingehend untersucht hat, bezeichnet die fragliche Operation als *conjonction*. Dem Prozeß begegnet man schon bei C. Jordan, *Traité des substitutions et des équations algébriques*, Paris 1870, S. 221. Die Determinante von $A_1 \times A_2$ hat Kronecker in seinen Universitätsvorlesungen behandelt. Von der Operation $A \times A \times \dots \times A$ (m mal die nämliche Matrix A) kann man zu $P_m(A)$ und $C_m(A)$ gelangen (vgl. J. Schur und A. Loewy an den S. 146 angeführten Orten).

Für die Produkttransformation gelten folgende Sätze:

I. Sind $B_1, B_2, \dots, B_r$ r weitere Matrices der Grade $n_1, n_2, \dots, n_r$, aus denen $B_1 \times B_2 \times \dots \times B_r$ gebildet ist, so ist das Produkt $W_1 W_2$ von $W_1 = A_1 \times A_2 \times \dots \times A_r$ und $W_2 = B_1 \times B_2 \times \dots \times B_r$ gleich $A_1 B_1 \times A_2 B_2 \times \dots \times A_r B_r$.

II. Sind $A_1, A_2, \dots, A_r$ Einheitsmatrices, so ist auch $A_1 \times A_2 \times \dots \times A_r$ die Einheitsmatrix.

III. Die Matrix der Produkttransformation der reziproken Matrices von $A_1, A_2, \dots, A_r$ ist die reziproke Matrix von $A_1 \times A_2 \times \dots \times A_r$, also

$$A_1^{-1} \times A_2^{-1} \times \dots \times A_r^{-1} = (A_1 \times A_2 \times \dots \times A_r)^{-1}.$$

IV. Die transponierte Matrix von $A_1 \times A_2 \times \dots \times A_r$ ist die Matrix der Produkttransformation der transponierten

Matrices $A_1', A_2', \ldots, A_r'$, also:

$$(A_1 \times A_2 \times \cdots \times A_r)' = (A_1' \times A_2' \times \cdots \times A_r').$$

V. Sind $\alpha_1^{(\nu)}, \alpha_2^{(\nu)}, \ldots, \alpha_{n_\nu}^{(\nu)}$ die n_ν Wurzeln der charakteristischen Gleichung der Matrix $A^{(\nu)}$, d. h. der Gleichung $|\varrho E - A^{(\nu)}| = 0$, so sind die $n_1 n_2 \ldots n_r$ Wurzeln der charakteristischen Gleichung von $A_1 \times A_2 \times \cdots \times A_r$ die $n_1 n_2 \ldots n_r$ Produkte

$$\alpha_{i_1}^{(1)} \alpha_{i_2}^{(2)} \ldots \alpha_{i_r}^{(r)} \quad (i_1 = 1, 2, \ldots, n_1;\ i_2 = 1, 2, \ldots, n_2;\ \ldots,\ i_r = 1, 2, \ldots, n_r)$$

(Franklin, *Am. J. math.* **16**, 206 (1894), Frobenius, *Sitzungsb. d. Berl. Akad.* (1899), 333, C. Stephanos, *Journ. de math.* (5) **6**, 89 (1900)).

Aus dem Lehrsatz V folgt: Die Determinante von

$$A_1 \times A_2 \times \cdots \times A_r$$

ist ein Produkt, dessen Faktoren die $\frac{\pi}{n_1}$te Potenz der Determinante von A_1, die $\frac{\pi}{n_2}$te Potenz der Determinante von A_2 usw., schließlich die $\frac{\pi}{n_r}$te Potenz der Determinante von A_r sind, also:

$$|A_1 \times A_2 \times \cdots \times A_r| = |A_1|^{\frac{\pi}{n_1}} |A_2|^{\frac{\pi}{n_2}} \ldots |A_r|^{\frac{\pi}{n_r}}, \text{ wobei } \pi = n_1 n_2 \ldots n_r$$

(Lehrsatz von Kronecker, vgl. Hensel, *Acta math.* **14**, 317 (1890), Rados, *Math. naturw. Ber. aus Ungarn* **8**, 60 (1889), ebenda **18**, 231 (1900), G. v. Escherich, *Monatsh. f. Math.* **3**, 68 (1892), A. Hurwitz, *Math. Ann.* **45**, 389 (1894)).

Zur Produkttransformation kann man auch auf folgende Weise gelangen:

$$H = \sum_{i_1 i_2 \ldots i_r} \eta_{i_1 i_2 \ldots i_r} x_{i_1}^{(1)} x_{i_2}^{(2)} \ldots x_{i_r}^{(r)}$$

sei eine Form, die in den r Variablenreihen

$$x_{i_1}^{(1)} \ (i_1 = 1, 2, \ldots, n_1),\quad x_{i_2}^{(2)} \ (i_2 = 1, 2, \ldots, n_2),\quad \ldots,\quad x_{i_r}^{(r)} \ (i_r = 1, 2, \ldots, n_r)$$

linear und homogen ist. Transformiert man die angegebene Form durch die linearen homogenen Substitutionen $A_1, A_2, \ldots, A_r$ in die neue r-fach lineare homogene Form:

$$Z = \sum_{i_1 i_2 \dots i_r} \zeta_{i_1 i_2 \dots i_r} \xi^{(1)}_{i_1} \xi^{(2)}_{i_2} \dots \xi^{(r)}_{i_r},$$

so wird:

$$(2) \qquad \zeta_{k_1 k_2 \dots k_r} = \sum_{i_1 i_2 \dots i_r} a^{(1)}_{i_1 k_1} a^{(2)}_{i_2 k_2} \dots a^{(r)}_{i_r k_r} \eta_{i_1 i_2 \dots i_r},$$

d. h. die $n_1 n_2 \dots n_r$ Koeffizienten $\eta_{i_1 i_2 \dots i_r}$ erleiden die Produkt transformation $A'_1 \times A'_2 \times \dots \times A'_r$; hierbei bedeuten $A'_1, A'_2, \dots, A'_r$ die transponierten Matrices von $A_1, A_2, \dots, A_r$.

Nimmt man in den r Variablenreihen $x_i^{(1)}, x_i^{(2)}, \dots, x_i^{(r)}$ gleichviele Variablen n an und beschränkt in der Form H die n^r Größen

$$\eta_{i_1 i_2 \dots i_r} \quad (i_1 = 1, 2, \dots, n;\ i_2 = 1, 2, \dots, n;\ \dots,\ i_r = 1, 2, \dots, n),$$

symmetrische Funktionen der Indices zu sein, d. h. ihre Werte nicht zu ändern, wenn man die Reihenfolge der Indices vertauscht, so reduzieren sich die n^r Größen $\eta_{i_1 i_2 \dots i_r}$ auf $\binom{n+r-1}{r}$ verschiedene. Unterwirft man die r Variablenreihen von H alsdann r kogredienten linearen homogenen Substitutionen mit der nämlichen Matrix A:

$$x_i^{(\nu)} = a_{i1} \xi_1^{(\nu)} + a_{i2} \xi_2^{(\nu)} + \dots + a_{in} \xi_n^{(\nu)}$$
$$(i = 1, 2, \dots, n;\ \nu = 1, 2, \dots, r),$$

so bezieht sich bei den eingeführten Beschränkungen die durch die Formel (2) definierte lineare homogene Substitution auf $\binom{n+r-1}{r}$ Variablen und wird die r^{te} Potenztransformation von A', wobei A' die transponierte Matrix von A ist.

Wir wenden die obigen Resultate auf die bilineare Form:

$$H = \sum_{i=1}^{i=n} \sum_{k=1}^{k=n} \eta_{ik} x_i^{(1)} x_k^{(2)}$$

an, deren Koeffizienten η_{ik} zunächst n^2 willkürliche Größen seien. Unterwirft man die zwei Variablenreihen den zwei linearen homogenen Substitutionen:

$$A': \qquad x_i^{(1)} = a_{1i} \xi_1^{(1)} + a_{2i} \xi_2^{(1)} + \dots + a_{ni} \xi_n^{(1)}$$

und

$$B: \qquad x_i^{(2)} = b_{i1} \xi_1^{(2)} + b_{i2} \xi_2^{(2)} + \dots + b_{in} \xi_n^{(2)} \quad (i = 1, 2, \dots, n),$$

so geht H in eine neue bilineare Form

$$Z = \sum_{i=1}^{i=n} \sum_{k=1}^{k=n} \zeta_{ik} \xi_i^{(1)} \xi_k^{(2)}$$

über. Es gilt die symbolische Gleichung $Z = AHB$; die Determinante $|Z|$ ist gleich $|A||B||H|$. Die Koeffizienten ζ_{ik} ergeben sich gleich

$$\sum_{s=1}^{s=n} \sum_{t=1}^{t=n} a_{is} \eta_{st} b_{tk},$$

d. h. sie werden aus den η_{ik} durch Anwendung der Produkttransformation $A \times B'$ gefunden. Wird H so spezialisiert, daß es eine symmetrische bilineare Form wird, deren $\frac{n(n+1)}{2}$ Koeffizienten η_{ik} $(k \geqq i)$ beliebige Größen sind, und unterwirft man die zwei Variablenreihen $x^{(1)}$ und $x^{(2)}$ kogredienten Transformationen, d. h. wählt man $B = A'$, so ist $Z = AHA'$, und die Koeffizienten ζ_{ik} werden aus den η_{ik} durch Anwendung der zweiten Potenztransformation $P_2(A)$ von A gewonnen.

Mit diesen Resultaten stehen folgende zwei Fundamentalsätze in innigstem Zusammenhang:

I. Sind die Elemente der Matrix $H = \|\eta_{ik}\|$ $(i, k = 1, 2, \ldots, n)$ *unabhängige Variable und die der Matrix* $Z = \|\zeta_{ik}\|$ $(i, k = 1, 2, \ldots, n)$ *lineare homogene Funktionen dieser Variablen und unterscheidet sich die Determinante der Matrix Z von der der Matrix H nur um einen konstanten, von Null verschiedenen Faktor, so ist entweder* $Z = AHB$ *oder* $Z = AH'B$, *wobei A und B konstante Matrices sind* (Frobenius, *Sitzungsb. d. Berl. Akad.* (1897), 1011, S. Kantor, *Sitzungsb. d. Bayer. Akad.* (1897), 370, C. Stephanos, *Journ. d. math.* (5) **6**, 119ff. (1900), E. Steinitz, *Sitzungsb. d. Berl. Math. Gesellsch., Archiv f. Math.* (3) **5**, 47 (1903)).

II. Sind in einer symmetrischen Matrix $H = \|\eta_{ik}\|$ $(i, k = 1, 2, \ldots, n)$ *die Elemente* η_{ik} $(k \geqq i)$ *unabhängige Variable, und sind die Elemente der symmetrischen Matrix* $Z = \|\zeta_{ik}\|$ $(i, k = 1, 2, \ldots, n)$ *lineare homogene Funktionen dieser Variablen und unterscheidet sich die Determinante der Matrix Z von der der Matrix H nur um einen konstanten, von Null verschiedenen Faktor, so ist* $Z = AHA'$, *wobei A eine konstante Matrix bedeutet* (Frobenius, *Sitzungsb. d. Berl. Akad.* (1897), 1014).

Für $n = 2$ werden durch das Theorem I alle linearen homogenen Substitutionen bestimmt, die

$$|H| = \begin{vmatrix} \eta_{11} & \eta_{12} \\ \eta_{21} & \eta_{22} \end{vmatrix},$$

abgesehen von einem konstanten Faktor, automorph in $\xi_{11}\xi_{22} - \xi_{12}\xi_{21}$ transformieren. Mit diesem Formelsystem hat sich Cayley bereits 1854 beschäftigt, *Phil. Mag., Coll. math. papers* **2**, 135 „On the homographic transformation of a surface of the second order into itself". Vgl. Fricke und Klein, *Vorl. über die Theorie der automorphen Funktionen*, Leipzig 1897, S. 47.

Der Satz II bestimmt für $n = 2$ alle Transformationen, die

$$\begin{vmatrix} \eta_{11} & \eta_{12} \\ \eta_{12} & \eta_{22} \end{vmatrix},$$

abgesehen von einem konstanten Faktor, in $\xi_{11}\xi_{22} - \xi_{12}^2$ transformieren. Dieses Formelsystem, das den Kegelschnitt

$$\eta_{11}\eta_{22} - \eta_{12}^2 = 0$$

in sich überführt, hat bereits Gauß, *Disquisitiones arithmeticae*, Artikel 157, *Ges. Werke* **1**, 124; vgl. Fricke und Klein, a. a. O., S. 14 und 19.

§ 12. Differentiation einer Determinante und Matrix.

Sind y_{ik} $(i,k = 1,2,\ldots,n)$ n^2 Funktionen einer Variablen x, so gilt der Satz:

Der Differentialquotient der Determinante $|y_{ik}|$ *ist gleich der Summe aller n Determinanten, die man erhält, wenn man in der gegebenen Determinante immer die Elemente einer Reihe durch ihre Differentialquotienten ersetzt.*

Versteht man unter Y die Matrix $\|y_{ik}\|$, so wird unter $\frac{dY}{dx}$ die Matrix $\left\|\frac{dy_{ik}}{dx}\right\|$ verstanden. Sind Y und Z zwei Matrices gleichen Grades, deren Koeffizienten von einer Variablen x abhängen, so wird, wenn YZ das Produkt der zwei Matrices Y und Z ist,

$$\frac{d}{dx}(YZ) = \frac{dY}{dx}Z + Y\frac{dZ}{dx}, \quad \frac{d}{dx}Y^{-1} = -Y^{-1}\frac{dY}{dx}Y$$

(vgl. Frobenius, *Journ. f. Math.* **84**, 16 (1878)).

Das Produkt der zwei Matrices Y^{-1} und $\frac{dY}{dx}$, also $Y^{-1}\frac{dY}{dx}$, bezeichnet man mit $D_x(y_{ik})$ und liest es *derivierte Matrix in*

bezug auf x. Das eingeführte Symbol ist für die Systeme linearer homogener Differentialgleichungen von größter Wichtigkeit (Volterra, *Memorie della società Italiana delle scienze* (1887), (1899), L. Schlesinger, *Vorl. über die Theorie der linearen Differentialgleichungen,* Leipzig 1908, S. 26).

§ 13. Die Wronskische Determinante.

Hat man n Funktionen $y_1, y_2, \ldots, y_n$ einer Variablen x, so entsteht ihre *Wronskische Determinante* (Wronski, *Réfutation de la théorie des fonctions analytiques de Lagrange,* 1812, Dickstein, *Bibliotheca math.* (2) **6**, 50 (1892), E. Pascal, *Acc. di Torino* (1906), 1081) auf folgende Art:

In die erste Zeile kommen die n Funktionen, in die folgenden die ersten, zweiten, usw. Abgeleiteten dieser Funktionen; also:

$$W = \begin{vmatrix} y_1 & y_2 & \cdots & y_n \\ \frac{dy_1}{dx} & \frac{dy_2}{dx} & \cdots & \frac{dy_n}{dx} \\ \frac{d^2y_1}{dx^2} & \frac{d^2y_2}{dx^2} & \cdots & \frac{d^2y_n}{dx^2} \\ \vdots & \vdots & & \vdots \\ \frac{d^{n-1}y_1}{dx^{n-1}} & \frac{d^{n-1}y_2}{dx^{n-1}} & \cdots & \frac{d^{n-1}y_n}{dx^{n-1}} \end{vmatrix}.$$

Die Abgeleitete einer Wronskischen Determinante wird gebildet, indem man in die letzte Zeile von W die n^{ten} Abgeleiteten setzt und die anderen Zeilen ungeändert läßt (Malmsten, *Journ. f. Math.* **39**, 93 (1850)).

Multipliziert man die n Funktionen y mit einer beliebigen Funktion $v(x)$, so wird hierdurch die ganze Determinante mit der n^{ten} Potenz von v multipliziert.

Die notwendige und auch im allgemeinen hinreichende Bedingung dafür, daß zwischen n Funktionen $y_1(x), y_2(x), \ldots, y_n(x)$ wenigstens eine lineare homogene Beziehung mit konstanten Koeffizienten besteht, ist das Verschwinden ihrer Wronskischen Determinante (vgl. besonders Frobenius, *Journ. f. Math.* **77**, 245 (1874)).

Peano (*Mathesis*, **9**, 75 u. 110 (1889), sowie *Rend. Acc. Lincei* 6_1, 413 (1897)) hat zuerst darauf aufmerksam gemacht, daß das Verschwinden der Wronskischen Determinante von n Funktionen einer Variablen nicht unter allen Umständen für

ihre lineare Dependenz ausreicht; für analytische Funktionen ist die Bedingung hinreichend. Vgl. auch die Angaben von Bôcher, *Trans. Am. M. S.* **2**, 139 (1901), über gewisse Fälle nicht analytischer Funktionen, bei denen das Verschwinden der Wronskischen Determinante für die lineare Abhängigkeit ausreicht, ferner D. R. Curtiss, *Math. Ann.* **65**, 282 (1908).

Die Wronskische Determinante spielt in der Theorie der linearen homogenen Differentialgleichungen eine sehr wichtige Rolle (vgl. Baltzer, *Determinanten*, S. 77, L. Heffter, *Einl. in die Theorie der linearen Differentialgleichungen*, Leipzig 1894, S. 47 u. 233, L. Schlesinger, *Handbuch der Theorie der linearen Differentialgleichungen*, Leipzig 1895, **1**, S. IX u. 37).

Eine ähnliche Bedeutung wie die Wronskische Determinante für die linearen homogenen Differentialgleichungen hat für die Theorie der *linearen homogenen Differenzengleichungen* die Determinante:

$$W_1 = \begin{vmatrix} y_1(x) & y_2(x) \dots & y_n(x) \\ \Delta y_1(x) & \Delta y_2(x) \dots & \Delta y_n(x) \\ \vdots & \vdots & \vdots \\ \Delta^{n-1} y_1(x) & \Delta^{n-1} y_2(x) \dots & \Delta^{n-1} y_n(x) \end{vmatrix};$$

hierbei ist Δ das Symbol für die Differenzen:

$$\begin{aligned} \Delta y(x) &= y(x+1) - y(x), \\ \Delta^2 y(x) &= \Delta y(x+1) - \Delta y(x), \\ &\vdots \end{aligned}$$

W_1 ist gleich der Determinante:

$$\begin{vmatrix} y_1(x) & y_2(x) & \dots y_n(x) \\ y_1(x+1) & y_2(x+1) & \dots y_n(x+1) \\ y_1(x+2) & y_2(x+2) & \dots y_n(x+2) \\ \vdots & \vdots & \vdots \\ y_1(x+n-1) & y_2(x+n-1) & \dots y_n(x+n-1) \end{vmatrix}.$$

Das Verschwinden der Determinante W_1 ist die notwendige und hinreichende Bedingung dafür, daß zwischen den n Funktionen y eine lineare homogene Beziehung mit Koeffizienten besteht, die periodische Funktionen von x sind, d. h. solche Funktionen $F(x)$, bei denen für jedes beliebige x:

$$F(x+1) = F(x)$$

ist (Casorati, *Ann. mat.* (2) **10**, 19 (1880)).

In der Determinante W_1 kann statt Δ irgendeine Funktionaloperation und ihre Wiederholungen eingeführt werden, derartige Determinanten bei Pincherle u. Amaldi, *Le operazioni distributive,* Bologna 1901, 419 u. 429.

§ 14. Die Jacobische oder Funktionaldeterminante.

Sind n Funktionen $y_1, y_2, \ldots, y_n$ von n Variablen $x_1, x_2, \ldots, x_n$ gegeben, so heißt die Determinante:

$$\begin{vmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} & \cdots & \frac{\partial y_2}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial y_n}{\partial x_1} & \frac{\partial y_n}{\partial x_2} & \cdots & \frac{\partial y_n}{\partial x_n} \end{vmatrix}$$

die *Funktionaldeterminante* oder *Jacobische Determinante* der y in bezug auf die x. Man bezeichnet sie auch nach Donkin (*Phil. Trans.* (1854)) mit dem Symbol:

$$\frac{\partial (y_1 y_2 \ldots y_n)}{\partial (x_1 x_2 \ldots x_n)}.$$

Der Name „Funktionaldeterminante" stammt von Jacobi; er ist von ihm eingeführt in der grundlegenden Arbeit *„de determinantibus functionalibus"*, *Journ. f. Math.* **22**, 319 (1841), *Ges. Werke* **3**, 393, deutsche Ausg. mit Anm. von Stäckel in Ostwalds *Klassikern der exakten Wiss.* No. 78. Die Bezeichnung „Jacobische Determinante" geht auf Sylvester (*Phil. Trans.* (1853), *Coll. math. papers* **1**, 506 u. 583) zurück.

Sind $y_1, y_2, \ldots, y_n$ Funktionen von $z_1, z_2, \ldots, z_n$ und diese wieder Funktionen von $x_1, x_2, \ldots, x_n$, so gilt die Formel:

$$\frac{\partial(y_1 y_2 \dots y_n)}{\partial(x_1 x_2 \dots x_n)} = \frac{\partial(y_1 y_2 \dots y_n)}{\partial(z_1 z_2 \dots z_n)} \cdot \frac{\partial(z_1 z_2 \dots z_n)}{\partial(x_1 x_2 \dots x_n)}.$$

Sind $y_1, y_2, \dots, y_n$ Funktionen von $x_1, x_2, \dots, x_n$ und kann man $x_1, x_2, \dots, x_n$ als Funktionen von $y_1, y_2, \dots, y_n$ ansehen, so hat man:

$$\frac{\partial(y_1 y_2 \dots y_n)}{\partial(x_1 x_2 \dots x_n)} = \frac{1}{\dfrac{\partial(x_1 x_2 \dots x_n)}{\partial(y_1 y_2 \dots y_n)}}.$$

Sind die Größen y implizit als Funktionen der Veränderlichen x mittels der Gleichungen:

$$\begin{aligned} F_1(y_1, y_2, \dots, y_n, x_1, x_2, \dots, x_n) &= 0, \\ F_2(y_1, y_2, \dots, y_n, x_1, x_2, \dots, x_n) &= 0, \\ &\vdots \\ F_n(y_1, y_2, \dots, y_n, x_1, x_2, \dots, x_n) &= 0 \end{aligned}$$

gegeben, so ist:

$$\frac{\partial(y_1 y_2 \dots y_n)}{\partial(x_1 x_2 \dots x_n)} = (-1)^n \frac{\dfrac{\partial(F_1 F_2 \dots F_n)}{\partial(x_1 x_2 \dots x_n)}}{\dfrac{\partial(F_1 F_2 \dots F_n)}{\partial(y_1 y_2 \dots y_n)}}.$$

Das Nichtverschwinden der im Nenner stehenden Funktionaldeterminante

$$\frac{\partial(F_1 F_2 \dots F_n)}{\partial(y_1 y_2 \dots y_n)}$$

ist die Bedingung dafür, daß durch die Gleichungen

$$F_i(y_1, y_2, \dots, y_n, x_1, x_2, \dots, x_m) = 0 \quad (i = 1, 2, \dots, n)$$

ein System von n Funktionen $y_1, y_2, \dots, y_n$ der m Variablen $x_1, x_2, \dots, x_m$ definiert wird. Existenzsatz für die impliziten Funktionen (vgl. etwa C. Jordan, *Cours d'analyse,* Paris 1893, **1**, 82, Genocchi-Peano, *Differentialrechnung und Grundzüge der Integralrechnung,* deutsche Ausg., Leipzig 1899, S. 147).

Soll zwischen n Funktionen y_i der n Variablen $x_1, x_2, \dots, x_n$ eine von den x freie Beziehung: $F(y_1, y_2, \dots, y_n) = 0$ bestehen, so ist die notwendige und hinreichende Bedingung hierfür die, daß die Funktionaldeterminante der y verschwindet.

Schärfer und weitgehender ist folgendes Theorem:

Sind m Funktionen $y_1, y_2, \ldots, y_m$ von n Variablen $x_1, x_2, \ldots, x_n$ gegeben und verschwindet die Determinante l^{ten} Grades[1])

$$\frac{\partial(y_1\, y_2 \ldots y_l)}{\partial(x_1\, x_2 \ldots x_l)}$$

für ein besonderes Wertsystem der Variablen

$$x_1 = \xi_1,\ x_2 = \xi_2,\ \ldots,\ x_n = \xi_n$$

nicht, sind hingegen alle Determinanten $l+1^{ten}$ Grades, die man aus der Matrix

$$\left\| \frac{\partial y_i}{\partial x_k} \right\| \qquad (i = 1, 2, \ldots, m;\ k = 1, 2, \ldots, n)$$

bilden kann, in der Umgebung der Stelle $\xi_1, \xi_2, \ldots, \xi_n$ identisch Null, so sind die Größen $y_1, y_2, \ldots, y_l$ unabhängig, d. h. keine von ihnen läßt sich als Funktion der $l-1$ übrigen darstellen, hingegen können $y_{l+1}, y_{l+2}, \ldots, y_m$ als Funktionen von $y_1, y_2, \ldots, y_l$ ausgedrückt werden.

Ist

$$y_i(x_1, x_2, \ldots, x_n) = \frac{u_i(x_1, x_2, \ldots, x_n)}{u_0(x_1, x_2, \ldots, x_n)} \qquad (i = 1, 2, \ldots, n),$$

so hat man:

$$\frac{\partial(y_1\, y_2 \ldots y_n)}{\partial(x_1\, x_2 \ldots x_n)} = \frac{1}{{u_0}^{n+1}} \begin{vmatrix} u_0 & \frac{\partial u_0}{\partial x_1} & \frac{\partial u_0}{\partial x_2} & \cdots & \frac{\partial u_0}{\partial x_n} \\ u_1 & \frac{\partial u_1}{\partial x_1} & \frac{\partial u_1}{\partial x_2} & \cdots & \frac{\partial u_1}{\partial x_n} \\ \vdots & \vdots & \vdots & & \vdots \\ u_n & \frac{\partial u_n}{\partial x_1} & \frac{\partial u_n}{\partial x_2} & \cdots & \frac{\partial u_n}{\partial x_n} \end{vmatrix}.$$

Diese Theoreme sind sämtlich von Jacobi in *„de determinantibus functionalibus"* gegeben worden.

Es seien $n+1$ homogene ganze Funktionen von n Variablen gegeben. Man kombiniere sie zu je n und bilde so die $n+1$ Funktionaldeterminanten. Aus diesen bilde man wieder $n+1$ Funktionaldeterminanten, indem man sie zu n kombiniert. Die letzteren sind dann bis auf einen allen gemeinsamen Faktor die-

---

[1]) Durch geeignete Numerierung der y_i und x_k kann man es erzielen, daß gerade die Determinante

$$\frac{\partial(y_1\, y_2 \ldots y_l)}{\partial(x_1\, x_2 \ldots x_l)}$$

nicht verschwindet.

selben Funktionen wie diejenigen, von denen man ausgegangen ist (Theorem von Clebsch, *Journ. f. Math.* **69**, 355 (1868), **70**, 175 (1869)).

$y_1, y_2, \ldots, y_n$ seien n Funktionen von $n+1$ Variablen $x_1, x_2, \ldots, x_{n+1}$. Man betrachte die y als Funktionen von je n der Variablen x und bilde die Funktionaldeterminanten. Mit Vorzeichen versehen, seien sie mit A_i $(i = 1, 2, \ldots, n+1)$ bezeichnet, so daß

$$A_i = (-1)^i \frac{\partial (y_1 y_2 \ldots \ldots \ldots \ldots y_n)}{\partial (x_1 x_2 \ldots x_{i-1} x_{i+1} \ldots x_{n+1})}$$

ist, dann besteht die Jacobische Formel:

$$\frac{\partial A_1}{\partial x_1} + \frac{\partial A_2}{\partial x_2} + \cdots + \frac{\partial A_{n+1}}{\partial x_{n+1}} = 0,$$

(*Journ. f. Math.* **27**, 203 (1844), *Ges. Werke* **4**, 323). Vgl. auch Jacobis *Vorlesungen über Dynamik,* herausgeg. von Clebsch, 1866, 13. Vorl. Die Jacobische Formel spielt in der Theorie des Jacobischen Multiplikators bei einem System gewöhnlicher Differentialgleichungen eine fundamentale Rolle.

Die Funktionaldeterminanten sind auch für die Transformation der vielfachen Integrale sehr wichtig. Zum Zweck der Transformation der n fachen Integrale hat Jacobi die Funktionaldeterminanten bereits 1833 im *Journ. f. Math.* **12**, 38, *Ges. Werke* **3**, 233 verwandt und ist dann im letzten Paragraphen der Abhandlung *de determin. funct.* auf diesen Gegenstand zurückgekommen. Von Lehrbüchern vgl. man hierzu: Kronecker, *Vorl. über die Theorie der einfachen und vielfachen Integrale,* herausg. von Netto, Leipzig 1894, 14. Vorl., C. Jordan, *Cours d'analyse,* Paris 1893, **1**, 138 ff.

Eine von Bertrand (*Journ. de math.* **16**, 212 (1851)) gegebene Definition der Funktionaldeterminante (vgl. *Encycl. des sc. math.* **1**, 123) ist *unrichtig* (Genocchi-Peano, *Differentialrechnung usw.,* deutsche Ausg., Leipzig 1899, S. 329).

Grassmann (*Ges. Werke* $\mathbf{1}_2$, 298, vgl. auch Scott, *The theory of determinants,* S. 171) hat die Funktionaldeterminanten mit Hilfe der Methoden der Ausdehnungslehre behandelt.

Von Lehrbüchern vgl. man über Funktionaldeterminanten noch Pascal, *Determinanten,* S. 222, Baltzer, *Determinanten,* S. 139 u. Gordan, *Vorl. über Invariantentheorie* **1**, 120.

§ 15. Die Hessesche Determinante.

Die Hessesche Determinante einer Funktion f von n Variablen $x_1, x_2, \ldots, x_n$ *ist die Jacobische Determinante der n partiellen Abgeleiteten*

$$\frac{\partial f}{\partial x_i} \qquad (i = 1, 2, \ldots, n)$$

von f. Die Hessesche Determinante wird demnach durch die zweiten Abgeleiteten von f gebildet.

$$H = \left| \frac{\partial^2 f}{\partial x_i \partial x_k} \right| \qquad (i, k = 1, 2, \ldots, n),$$

(Hesse, *Journ. f. Math.* **28**, 89 (1844), *Ges. Werke,* München 1897, 113 u. 692). Die heute allgemein übliche Bezeichnung „Hessesche Determinante" stammt von Sylvester ((1851) u. (1853), *Coll. math. papers* **1**, 191 u. 583).

Die Hessesche Determinante einer beliebigen quadratischen Form ist das Doppelte ihrer Diskriminante (vgl. S. 118).

Die notwendige Bedingung dafür, daß eine homogene ganze Funktion (Form) von n Veränderlichen sich linear homogen in eine solche von weniger als n Veränderlichen transformieren läßt, ist das Verschwinden ihrer Hesseschen Determinante.

Daß umgekehrt das Verschwinden der Hesseschen Determinante, wie Hesse annahm (*Journ. f. Math.* **42**, 117 (1851), ebenda **56**, 263 (1859), *Ges. Werke*, 289 u. 481), auch dafür ausreichend ist, daß sich die gegebene homogene Form linear in eine solche von weniger Variablen transformieren läßt, *trifft nur für homogene Formen beliebigen Grades in 2, 3 und 4 Variablen (binäre, ternäre, quaternäre Formen) und quadratische Formen von beliebig vielen Variablen zu. Für die höheren Fälle lassen sich ganze Klassen von Funktionen aufstellen, bei denen die Hessesche Determinante verschwindet, ohne daß sie sich in Funktionen von weniger Variablen linear transformieren lassen* (Gordan u. Noether, *Math. Ann.* **10**, 547 (1876)).

Von der Hesseschen Determinante einer homogenen ganzen Funktion dritten Grades mit beliebig vielen Variablen gilt folgender Satz:

Die Hessesche Determinante der Hesseschen Determinante einer kubischen Form von n Variablen ist eine lineare Kombination der gegebenen Form und ihrer Hesseschen Determinante.

Dieser Satz wurde für beliebiges n von Voss, *Math. Ann.* **27**, **515** (1886), vorher für $n = 4$ von G. Bauer, *Abh. d. Bayer.*

Akad. d. Wiss. **14** (1883), bewiesen. Für binäre Formen gilt der Satz für jeden Grad > 3 (Salmon-Fiedler, *Algebra d. linearen Transformationen,* Leipzig 1877, S. 273).

Die Hessesche Determinante spielt ebenso wie die Funktionaldeterminante in der Invariantentheorie sowie in der Geometrie der Kurven und Flächen eine wichtige Rolle.

§ 16. Unendliche Determinanten. Systeme von unendlich vielen linearen Gleichungen. Bilineare Formen mit unendlich vielen Variablen.

Unendliche Determinanten wurden von Hill, *Acta math.* **8**, 26 (1886), Poincaré, *Bull. de la soc. math.* **14**, 77 (1886) und Helge von Koch, *Acta math.* **15**, 56 (1891) u. **16**, 219 (1892) in die Analysis eingeführt. Eine zusammenfassende Darstellung gibt Cazzaniga (*Ann. di mat.* (2) **26**, 143 (1897), Ergänzungen (3) **1**, 83 (1898) und (3) **2**, 229 (1899)). Wie die gewöhnlichen Determinanten zur Auflösung eines Systems von endlich vielen linearen Gleichungen, so dienen die unendlichen Determinanten zur Auflösung eines Systems mit unendlich vielen Unbekannten.

Die unendlichen Determinanten werden nach Cazzaniga auf folgende Weise definiert: Die Größen a_{ik} $(i,k = -\infty, \ldots, 0, \ldots, +\infty)$ mögen ein zweifach unendliches System bilden, das quadratisch nach vier Seiten hin unbegrenzt angeordnet sei. $D_{m,n}$ bedeute die Determinante $m + n + 1^{\text{ten}}$ Grades:

$$D_{m,n} = \begin{vmatrix} a_{-n,-n} & a_{-n,-n+1} & \cdots & a_{-n,0} & \cdots & a_{-n,m} \\ a_{-n+1,-n} & a_{-n+1,-n+1} & \cdots & a_{-n+1,0} & \cdots & a_{-n+1,m} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{m,-n} & a_{m,-n+1} & \cdots & a_{m,0} & \cdots & a_{m,m} \end{vmatrix}$$

Nähert sich für unendlich wachsende Werte von m und n die Determinante $D_{m,n}$ einer bestimmten endlichen Größe, so schreibt man

$$D = [a_{ik}] \qquad (i,k = -\infty, \ldots, 0, \ldots, +\infty)$$

und sagt: *die unendliche Determinante ist konvergent und hat den Wert D.*

Damit $\lim_{m=\infty,\, n=\infty} D_{m,n} = D$ existiert, also das Größensystem a_{ik} eine konvergente Determinante erzeugt, ist notwendig und hinreichend, daß zu jeder noch so kleinen positiven Zahl δ eine ganze positive Zahl N existiert, daß für jedes Zahlenpaar $m, n > N$ und alle ganzzahligen positiven Werte p, q:

$$|D_{m+p,n+q} - D_{m,n}| < \delta$$

wird.

Die Elemente a_{ii} $(i = -\infty, \ldots, 0, \ldots, +\infty)$ heißen die *Diagonalelemente,* die Elemente a_{ik} $(i \gtrless k)$ die *Nichtdiagonalelemente* der unendlichen Determinante.

Eine auf die obige Weise definierte unendliche Determinante hat ähnliche Eigenschaften wie die endlichen Determinanten. Der Wert der Determinante bleibt ungeändert, wenn man sie umstürzt, d. h. die Zeilen und die Kolonnen miteinander vertauscht und hierbei die Hauptdiagonale beibehält. Bei Vertauschung zweier paralleler Reihen ändert die Determinante ihr Vorzeichen; sind also zwei Parallelreihen identisch, so verschwindet die Determinante.

Jede unendliche konvergente Determinante $[a_{ik}]$ $(i,k = -\infty, \ldots, 0, \ldots, +\infty)$ *kann in eine gleichwertige transformiert werden, deren quadratisch geordnete Elemente nur nach rechts und nach unten unbegrenzt fortschreiten, d. h. in eine Determinante* $[a_{\lambda\mu}]$ $(\lambda, \mu = 1, 2, \ldots, \infty)$.

Dies erzielt man, indem man die Zeilen und die Kolonnen der ursprünglichen Determinante etwa derartig vertauscht, daß $a_{00}\, a_{11}\, a_{-1,-1}\, a_{22}\, a_{-2,-2}\, a_{33}\, a_{-3,-3} \ldots$ zum Diagonalglied wird und hierauf die Zahlenreihe 0, 1, — 1, 2, —2, . . . mit 1, 2, 3, . . . bezeichnet.

Geht man umgekehrt von einer konvergenten Determinante[1] $[a_{\lambda\mu}]$ $(\lambda, \mu = 1, 2, \ldots, \infty)$ aus und ändert diese außerdem bei jeder Vertauschung ihrer Zeilen und Kolonnen, bei der die Diagonalglieder in der Hauptdiagonale stehen bleiben, nach Voraussetzung ihren Wert nicht, so kann man sie in eine gleichwertige Determinante $[a_{ik}]$ $(i,k = -\infty, \ldots, 0, \ldots, +\infty)$ verwandeln, deren

---

1) Eine unendliche Determinante $[a_{\lambda\mu}]$ $(\lambda, \mu = 1, 2, \ldots, \infty)$, deren Elemente nur nach zwei Seiten hin unbegrenzt ausgedehnt sind, heißt konvergent, wenn sich die Determinante $D_n = |a_{\lambda\mu}|$ $(\lambda, \mu = 1, 2, \ldots, n)$ für unendliche n einer bestimmten endlichen Grenze nähert.

quadratisches Schema die ganze Ebene bedeckt. Man setze $i = -\frac{\lambda-1}{2}$ für ungerade λ und $i = \frac{\lambda}{2}$ für gerade λ, ebenso $k = -\frac{\mu-1}{2}$ für ungerade μ und $k = \frac{\mu}{2}$ für gerade μ und führe die notwendigen Zeilen- und Kolonnenvertauschungen aus, so daß $\ldots a_{-2,-2}\, a_{-1,-1}\, a_{00}\, a_{11}\, a_{22} \ldots$ Diagonalglied wird.

Man kann sich daher auf die Betrachtung einer in dem obigen Sinn konvergenten Determinante $[a_{\lambda\mu}]$ $(\lambda, \mu = 1, 2, \ldots, \infty)$ beschränken. Im folgenden wird dies geschehen.

Sind die unendlichen Produkte $\prod_{\lambda=1}^{\lambda=\infty} p_\lambda = p$ *und* $\prod_{\mu=1}^{\mu=\infty} q_\mu = q$ *unbedingt konvergent und nicht Null, so hat die aus der Determinante* $D = [a_{\lambda\mu}]$ $(\lambda, \mu = 1, 2, \ldots, \infty)$ *dadurch hervorgehende Determinante, daß man alle Elemente der* λ^{ten} *Zeile mit* p_λ $(\lambda = 1, 2, \ldots, \infty)$ *und alle Elemente der* μ^{ten} *Kolonne mit* q_μ $(\mu = 1, 2, \ldots, \infty)$ *multipliziert, den Wert* pqD.

Eine besondere Gattung im obigen Sinn konvergenter Determinanten sind diejenigen, bei denen die unendliche Doppelreihe der Nichtdiagonalglieder $\sum_{\lambda=1}^{\lambda=\infty}\sum_{\mu=1}^{\mu=\infty} a_{\lambda\mu}$ $(\lambda \gtrless \mu)$ *und das unendliche Produkt der Diagonalglieder* $\prod_{\lambda=1}^{\lambda=\infty} a_{\lambda\lambda}$ *unbedingt konvergieren.* Unendliche Determinanten, die diesen Bedingungen genügen, heißen nach von Koch *Normaldeterminanten* (*Acta math.* **16**, 221).

Streicht man in einer Normaldeterminante r Zeilen und r Kolonnen, so erhält man wiederum eine *Normaldeterminante*. Ebenso wie bei endlichen Determinanten kann man bei Normaldeterminanten den Begriff der *Unterdeterminante* einführen. Versteht man unter $A_{\varrho\sigma}$ die mit $(-1)^{\varrho+\sigma}$ multiplizierte Determinante, die aus einer Normaldeterminante durch Streichen der ϱ^{ten} Zeile und σ^{ten} Kolonne hervorgeht, so gilt in Erweiterung der bei endlichen Determinanten stattfindenden Relationen (S. 58):

$$a_{\varrho 1}A_{\sigma 1} + a_{\varrho 2}A_{\sigma 2} + a_{\varrho 3}A_{\sigma 3} + \cdots = \delta_{\varrho\sigma}D \quad (\varrho, \sigma = 1, 2, \ldots),$$

$$a_{1\varrho}A_{1\sigma} + a_{2\varrho}A_{2\sigma} + a_{3\varrho}A_{3\sigma} + \cdots = \delta_{\varrho\sigma}D \quad (\varrho, \sigma = 1, 2, \ldots).$$

Auch der Laplacesche Zerlegungssatz (S. 58) ist erweiterungsfähig.

Sind $A = [a_{\lambda\mu}]$ $(\lambda, \mu = 1, 2, \ldots, \infty)$ und $B = [b_{\lambda\mu}]$ $(\lambda, \mu = 1, 2, \ldots, \infty)$

zwei unendliche Normaldeterminanten und bezeichnet man mit $c_{\lambda\mu}$ eine der vier unendlichen Reihen

$$\sum_{s=1}^{s=\infty} a_{\lambda s} b_{s\mu}, \quad \sum_{s=1}^{s=\infty} a_{\lambda s} b_{\mu s}, \quad \sum_{s=1}^{s=\infty} a_{s\lambda} b_{\mu s}, \quad \sum_{s=1}^{s=\infty} a_{s\lambda} b_{s\mu},$$

so ist jede der vier Determinanten $C = [c_{\lambda\mu}]$ $(\lambda, \mu = 1, 2, \ldots, \infty)$ ebenfalls eine Normaldeterminante und das Produkt von A und B (*Erweiterung des Cauchy-Binetschen Multiplikationssatzes* (S. 58)).

Sei das System unendlich vieler linearer Gleichungen

$$a_{\lambda 1} x_1 + a_{\lambda 2} x_2 + \cdots = y_\lambda \qquad (\lambda = 1, 2, \ldots, \infty)$$

mit normaler Determinante $D = [a_{\lambda\mu}]$ $(\lambda, \mu = 1, 2, \ldots, \infty)$ *gegeben.* Für $D \neq 0$ gibt es ein eindeutig bestimmtes Lösungssystem, bei dem die absoluten Beträge der Elemente x_λ $(\lambda = 1, 2, \ldots)$ sämtlich kleiner als eine angebbare positive endliche Zahl G sind. Es wird entsprechend den Cramerschen Formeln, S. 74:

$$x_\lambda = \frac{\sum_{\mu=1}^{\mu=\infty} A_{\mu\lambda} y_\mu}{D}.$$

Das homogene System

$$y_\lambda = 0 \qquad (\lambda = 1, 2, \ldots, \infty)$$

hat, falls die Determinante D nicht verschwindet, nur die triviale Lösung $x_\lambda = 0$ $(\lambda = 1, 2, \ldots, \infty)$.

Außer den Normaldeterminanten sind noch weitere unendliche konvergente Determinanten untersucht worden (vgl. außer der bereits zitierten Literatur noch H. v. Koch, *Acta math.* **24**, 89 (1901)).

Über die Lösung linearer Gleichungen mit unendlich vielen Unbekannten sind noch neuere Untersuchungen von Hilbert (*Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen,* 4. und 5. Mitteilung, *Gött. Nachr.* (1906), 157 u. 439, *Rend. Circolo di Palermo* **27** (1909) und E. Schmidt, *Rend. Circolo di Palermo* **25**, 53 (1908)) zu nennen.

In ihnen werden unendlich viele lineare Gleichungen

$$a_{\lambda 1} x_1 + a_{\lambda 2} x_2 + \cdots \qquad = y_\lambda \qquad (\lambda = 1, 2, \ldots, \infty)$$

untersucht; hierbei wird vorausgesetzt, daß in jeder einzelnen

Gleichung die Quadratsumme der absoluten Beträge der Koeffizienten, also

$$\sum_{\mu=1}^{\mu=\infty} |a_{\lambda\mu}|^2 \qquad (\lambda = 1, 2, \ldots, \infty)$$

konvergiert. Betrachtet werden nur *reguläre Lösungen,* d. h. solche, bei denen

$$\sum_{\lambda=1}^{\lambda=\infty} |x_\lambda|^2$$

konvergiert. Unter den für die Gleichungskoeffizienten gemachten Annahmen werden die notwendigen und hinreichenden Bedingungen für die Existenz regulärer Lösungen abgeleitet und alle solche Lösungen durch immer konvergente Reihen dargestellt. Ein durch diese Methoden lösbares unendliches Gleichungssystem braucht durchaus nicht eine konvergente unendliche Determinante zu haben, wie z. B. das einfache Beispiel:

$$\lambda x_\lambda = y_\lambda \qquad (\lambda = 1, 2, 3, \ldots)$$

lehrt. Ein Gleichungssystem

$$a_{\lambda 1} x_1 + a_{\lambda 2} x_2 + \cdots = y_\lambda \qquad (\lambda = 1, 2, \ldots)$$

mit Normaldeterminante $[a_{\lambda\mu}]$ erfüllt die Hilbert-Schmidtsche Bedingung für die Gleichungskoeffizienten; denn aus der unbedingten Konvergenz des unendlichen Produktes $\prod_{\lambda=1}^{\lambda=\infty} a_{\lambda\lambda}$ und der unendlichen Doppelreihe der Nichtdiagonalglieder folgt die Konvergenz von $\sum_{\mu=1}^{\mu=\infty} |a_{\lambda\mu}|^2$ $(\lambda = 1, 2, \ldots, \infty)$.

Die Behandlung eines Gleichungssystems mit Normaldeterminante beschränkt sich jedoch nicht, wie dies die Hilbert-Schmidtsche Methode tut, auf die regulären Lösungen; so hat das triviale System $x_\lambda = 1$ $(\lambda = 1, 2, 3, \ldots)$ eine Normaldeterminante, seine Lösungen lassen sich mittelst der erweiterten Cramerschen Formeln (S. 164) finden, sind aber nicht regulär.

Hilbert betrachtet auch quadratische und bilineare Formen und orthogonale Substitutionen in unendlich vielen Veränderlichen. Wir heben den Begriff der *beschränkten Bilinearform von unendlich vielen Variablen* hervor. *Eine bi-*

lineare Form von unendlich vielen Variablen und mit beliebigen Koeffizienten

$$A = \sum_{i=1}^{i=\infty} \sum_{k=1}^{k=\infty} a_{ik} x_i y_k$$

heißt beschränkt, wenn eine positive Größe M existiert, so daß für jedes n der absolute Betrag des n^{ten} Abschnittes, d. h. die Größe

$$|A_n| = |\sum_{i=1}^{i=n} \sum_{k=1}^{k=n} a_{ik} x_i y_k|$$

kleiner als M bleibt für alle x_i, y_k, für die $\sum_{i=1}^{i=n} |x_i|^2 = 1$, $\sum_{k=1}^{k=n} |y_k|^2 = 1$.

Hat man zwei beschränkte Bilinearformen A und B von unendlich vielen Variablen, so existiert ihr Produkt AB genau wie bei Bilinearformen von endlich vielen Variablen (vgl. § 6); das Produkt AB ist wiederum eine beschränkte Bilinearform. Für das Produkt dreier beschränkter bilinearer Formen gilt das assoziative Gesetz. Hingegen gibt es bei beschränkten Bilinearformen mit unendlich vielen Variablen nicht mehr notwendig eine einzige *reziproke* oder *inverse* beschränkte bilineare Form. Ist z. B. $A = \sum_{i=1}^{i=\infty} x_i y_{i+1}$, so ist $AX = E$ lösbar durch $\sum_{i=1}^{i=\infty} a_i x_1 y_i + x_{i+1} y_i$, wobei $a_1, a_2, \ldots$ willkürlich bleiben. Besitzt eine unendliche Bilinearform sowohl eine vordere als eine hintere Reziproke, so sind beide einander gleich und daher die einzigen (vgl. Hellinger u. Toeplitz, *Gött. Nachr.* (1906), Toeplitz, ebenda (1907), 101 u. 110).

Über die Verwendung unendlicher Determinanten in der Theorie der linearen Differentialgleichungen vgl. Schlesinger, *Handbuch der Theorie der linearen Differentialgleichungen* **1**, 274, $\mathbf{2}_1$, 529, Horn, *Gewöhnliche Differentialgleichungen, Samml. Schubert* **50**, Leipzig 1905, S. 206.

Wegen unendlicher Determinanten vgl. auch die Darstellung von A. Pringsheim in der *Encykl. d. math. Wiss.* **1**, 141.

§ 17. Kubische Determinanten und solche höheren Ranges.

Sind a_{fgh} n^3 Elemente mit drei Indices, so heißt die Summe von $(n!)^2$ Gliedern $\varepsilon a_{1g_1h_1} a_{2g_2h_2} \dots a_{ng_nh_n}$ eine *kubische Determinante* $D^{(1)}$, falls $g_1, g_2, \dots, g_n$ und $h_1, h_2, \dots, h_n$ jede Anordnung der Zahlen $1, 2, \dots, n$ bedeuten und ε die positive oder negative Einheit ist, je nachdem die Summe der Inversionen in den zwei Anordnungen $g_1, g_2, \dots, g_n$ und $h_1, h_2, \dots, h_n$ eine gerade oder ungerade Zahl ist. In einer kubischen Determinante sind $n!$ gewöhnliche Determinanten enthalten.

Anstatt die ersten Indices in der Reihenfolge $1, 2, \dots, n$ geordnet anzunehmen und die zweiten und dritten Indices zu permutieren, kann man auch die zweiten Indices natürlich geordnet denken, die ersten und dritten permutieren und nach ihnen die Vorzeichenbestimmung vornehmen. Dann erhält man eine *kubische Determinante* $D^{(2)}$. Durch Festbleiben der dritten Indices kann schließlich eine *kubische Determinante* $D^{(3)}$ definiert werden. Die drei kubischen Determinanten $D^{(1)}$, $D^{(2)}$ und $D^{(3)}$ sind im allgemeinen verschieden.

Kubische Determinanten wurden zuerst (1861) von de Gasparis behandelt. Wegen ihrer Eigenschaften und Literatur sei auf Pascal, *Determinanten,* S. 184, Günther, *Determinanten,* sowie Scott, *The theory of determinants,* S. 110, verwiesen; wir begnügen uns hier mit Hervorhebung von Hedrick, *Annals of math.* (2) **1**, 49 (1900), wegen der zahlreichen Literaturangaben.

Unendliche kubische Determinanten behandelt Cazzaniga, *Math. Ann.* **53**, 272 (1900).

Noch allgemeiner lassen sich *Determinanten höheren Ranges* als Summen von $(n!)^{\nu-1}$ Gliedern einführen; jedes Glied ist, abgesehen von der positiven oder negativen Einheit, ein Produkt von n Faktoren; jeder dieser weist ν Indices auf. Die Determinante $D^{(1)}$ wird aus dem Glied $a_{111\dots1} a_{222\dots2} \cdots a_{nnn\dots n}$ gewonnen, indem man bei den n Faktoren die ersten Indices festhält und die $\nu - 1$ folgenden auf jede Art anordnet; jedes Glied erhält das positive oder negative Vorzeichen, je nachdem die Summe der Inversionen, welche die $\nu - 1$ variablen Indices liefern, eine gerade oder ungerade Zahl ist. Wegen Determinanten höheren Ranges vgl. Gegenbauer, *Wiener Akad. Denkschriften* **43**, zweiter Teil, 17 (1882), **46**, 291 ff. (1883), G. v. Escherich, ebenda **43**, 1 (1882).

Kapitel III.

Algebraische Gruppentheorie.

Von *Alfred Loewy* in Freiburg i. Br.

§ 1. Historisches. Definition der Gruppe. Beispiele. Das Feld oder der Körper.

Mag auch die Idee der Gruppe „implizit so alt wie der mathematische Gedanke sein", so hat sich der Gruppenbegriff doch erst bei der Behandlung der algebraischen Gleichungen klar und deutlich herausgebildet. Lagrange wurde 1770 in seinen *Réflexions sur la résolution algébrique des équations, Œuvres* **3**, 205 (vgl. Pierpont, *Bull. Am. M. S.* **1**, 196 (1895)), Vandermonde in seiner *Résolution des équations* (1770) (deutsche Ausg., Berlin 1888) und vor allem Ruffini (vgl. Burkhardt, *Die Anfänge der Gruppentheorie und Paolo Ruffini, Ztschr. f. Math. u. Phys.* **37**, Suppl., S. 119 (1892) oder *Ann. di mat.* (2) **22**, 175 (1894)) bei seinen Unmöglichkeitsbeweisen für die algebraische Auflösbarkeit der Gleichungen von höherem als viertem Grad auf die Buchstabenvertauschungen oder die Lehre der *Permutationsgruppen* geführt. Cauchy (*J. éc. polyt.*, Cah. **17**, 1 u. 29 (1815), *Œuvres* (2) **1**, 64 u. 91, *Exercices d'analyse* **3** (1844), 151, *C. R.* **21** (1845), *Œuvres* (1) **9**, *C. R.* **22** (1846), *Œuvres* (1) **10**) hat die einzelnen Sätze seiner Vorgänger ergänzt, systematisch dargestellt und „in Terminologie und Bezeichnungsweise das Handwerkszeug geschaffen, dessen die Theorie der Permutationsgruppen zu ihrer Fortentwicklung benötigte". Auch Abel (*Journ. f. Math.* **1**, 65 (1826), **4**, 131 (1829), *Œuvres par Sylow et Lie* (1881), **1**, 28, 66 u. 478, **2**, 217 u. 329) ist zu nennen; denn wie er stets gewöhnt war, „den höchsten Standpunkt einzunehmen", so sind auch seine algebraischen Arbeiten von gruppentheoretischen Ideen durchtränkt.

Galois (1811—1832) (*Œuvres*, publ. par J. Liouville, *Journ. de math.* (1), **11**, 381 (1846), separat von Picard (1897)) verdankt man den für Gruppen im weitesten Sinn fundamentalen Begriff der *invarianten* Untergruppe sowie die hierauf gegründete Einteilung der Gruppen in *einfache* und *zusammengesetzte*. Von Galois stammt auch die Bezeichnung „Gruppe", Cauchy spricht von einem „système de substitutions conjuguées", eine Bezeichnung, die sich auch in J. A. Serrets *Algèbre supérieure* **2**, 3. éd., Paris 1866, findet. Die außerordentliche Bedeutung der Permutationsgruppen erhellt aus Galois' Nachweis, daß zu jeder algebraischen Gleichung eine Permutationsgruppe, die sogenannte Galoissche Gruppe, gehört; sie spiegelt alle Eigenschaften der betreffenden Gleichung wieder. C. Jordans *Traité des substitutions et des équations algébriques*, Paris 1870, bedeutet für die Theorie der Permutationsgruppen einen Markstein.

Die analytische Darstellung der Permutationsgruppen führte Galois (*Œuvres par Picard*, p. 27) auf eine auch unabhängig von den Buchstabenvertauschungen zu definierende Klasse endlicher Gruppen, die nach F. Klein (*Math. Ann.* **17**, 63 (1880)) sogenannten *Kongruenzgruppen*. Diese Schöpfung Galois', auf das innigste mit den sogenannten Galoisschen Imaginären oder dem Galoisschen Feld $GF[p^n]$ (vgl. unten) verknüpft, liefert uns eine ganze Reihe einfacher Gruppen. Unter den Gruppen endlicher Ordnungen sind die einfachen Gruppen die selteneren, aber auch die interessanteren; für eine große Anzahl algebraischer Gleichungen, welche der Funktionentheorie und Geometrie entspringen, sind sie von größter Wichtigkeit. Eine zusammenfassende Darstellung unserer Kenntnisse von den Kongruenzgruppen gibt das auch an eigenen Resultaten reiche Werk von L. E. Dickson, *Linear groups with an exposition of the Galois field theory*, *Teubners Samml.* **6**, Leipzig 1901. Neben ihm kommen für die älteren Untersuchungen in diesem Zweig der Gruppentheorie C. Jordans *Traité* sowie Klein-Fricke, *Theorie der elliptischen Modulfunktionen*, 2 Bde., Leipzig 1890/92, in Betracht.

Daß bei einer Gruppe nicht die Art der Darstellung der Elemente, sondern das Gesetz für ihre Kombination den springenden Punkt bildet, es sich also um Fragen eines allgemeineren und abstrakteren Gedankenkreises handelt, sprachen zuerst Cayley (1854) und (1878) (*Coll. math. papers* **2**, 123 und **10**, 401) und Kronecker (*Monatsb. d. Berl. Akad.* (1870), *Werke* **1**, 274) aus. Der Zweig der Gruppentheorie, der die

Gruppe unabhängig von der Darstellung ihrer Elemente studiert, wird als *abstrakte* oder *allgemeine Gruppentheorie* bezeichnet. Die Theorie der abstrakten Gruppen mit einer endlichen Anzahl von Elementen und die Theorie der Permutationsgruppen hängen auf das innigste durch den Satz zusammen: Jede endliche abstrakte Gruppe ist mit einer Permutationsgruppe holoedrisch isomorph, kann also durch eine Permutationsgruppe vertreten werden.

Der Begriff „Gruppe" ist aber nicht auf eine endliche Anzahl von Elementen beschränkt. Wäre man bei den linearen homogenen Differentialgleichungen genau auf die gleiche Weise vorgegangen wie Lagrange und Galois bei den algebraischen Gleichungen, so hätte man auf diesem Wege zu einem tieferen Studium der allgemeinen linearen homogenen Substitutionsgruppen und der Gruppen mit unendlich vielen Elementen gelangen können (vgl. E. Picard, *Traité d'analyse* **3**, Paris 1896, Chap. 16 u. 17, sowie die elementare Darstellung von A. Loewy, *Math. Ann.* **65**, 129 (1908)). Der historische Weg war ein anderer. Die Invariantentheorie der linearen homogenen Substitutionen oder, geometrisch gesprochen, die projektive Geometrie führte zu dem Begriff der Gruppe von unendlich vielen Elementen. Die Gruppe erscheint hier analytisch als die Gesamtheit der linearen homogenen Substitutionen:

$$x_i = a_{i1}x_1' + a_{i2}x_2' + \cdots + a_{in}x_n' \qquad (i = 1, 2, \ldots, n),$$

wobei die Determinante $|a_{ik}|$ $(i, k = 1, 2, \ldots, n)$ von Null verschieden ist. Deutet man die Variablen als homogene Punktkoordinaten eines $n-1$ dimensionalen Raumes, so hat man alle kollinearen Umformungen des R_{n-1}, d. h. alle Transformationen des R_{n-1}, die Punkte in Punkte überführen, vor sich. Die erste Arbeit, in der die unendliche Gruppe explizit eine wesentliche Rolle spielt, stammt von C. Jordan (*C. R.* **65**, 229 (1867), *Ann. di mat.* (2) **2**, 167 u. 322 (1869)). Bei Zugrundelegung rechtwinkliger Kartesischer Koordinaten findet eine Bewegung ihren analytischen Ausdruck in den Formeln:

$$\begin{aligned} x &= a_{11}x' + a_{12}y' + a_{13}z' + b_1, \\ y &= a_{21}x' + a_{22}y' + a_{23}z' + b_2, \\ z &= a_{31}x' + a_{32}y' + a_{33}z' + b_3, \end{aligned}$$

wobei die Matrix $\| a_{ik} \|$ $(i, k = 1, 2, 3)$ eine eigentliche orthogonale

Substitution (Kap. II, § 10) bestimmt. Die Gesamtheit aller angegebenen Operationen bildet die Bewegungsgruppe $\mathfrak{B}$. C. Jordans Problem ist das Studium aller Untergruppen der Bewegungsgruppe $\mathfrak{B}$.

Dann haben sich F. Klein und S. Lie des Gruppenbegriffes bemächtigt und vornehmlich ihnen ist es zu danken, daß sich die Gruppentheorie in fast allen Teilen der höheren Mathematik mehr und mehr Geltung verschaffte. S. Lie ist der Schöpfer der Theorie der kontinuierlichen Transformationsgruppen, die in einem besonderen Kapitel behandelt wird. Kleins sogenanntes Erlanger Programm (1872) *„Vergleichende Betrachtungen über neuere geometrische Forschungen"* (wiederabgedruckt: *Math. Ann.* **43**, 63 (1893)) will die ganze Geometrie als gruppentheoretisches Problem auffassen. Als weitere Ausführung dieser Ideen ist die autograph. Vorlesung *„Einleitung in die höhere Geometrie"* (Göttingen 1892/93) anzusehen. Vgl. auch Heffter und Köhler, *Lehrbuch der analytischen Geometrie,* Leipzig 1905. Bei Zugrundelegung rechtwinkliger Kartesischer Koordinaten erscheint die elementare Geometrie als das Studium aller Transformationen:

$$x = \lambda(a_{11}x' + a_{12}y' + a_{13}z') + b_1,$$
$$y = \lambda(a_{21}x' + a_{22}y' + a_{23}z') + b_2,$$
$$z = \lambda(a_{31}x' + a_{32}y' + a_{33}z') + b_3,$$

wobei die Matrix $\|a_{ik}\|$ $(i,k=1,2,3)$ eine eigentliche orthogonale Substitution bestimmt. Der Inbegriff aller dieser Transformationen hat Gruppencharakter, er setzt sich aus Bewegungen, die man erhält, wenn die willkürliche Konstante $\lambda = 1$ gewählt wird, Ähnlichkeitstransformationen, Spiegelungen sowie allen hieraus resultierenden Transformationen zusammen. Diese Gruppe heißt nach F. Klein die *Hauptgruppe.* Man vgl. für diese Fragen besonders den Artikel von Fano, *Kontinuierliche geometrische Gruppen. Die Gruppentheorie als geometrisches Einteilungsprinzip, Encykl. d. math. Wiss.* **3**, S. 289. Vorzüglich beschäftigte sich F. Klein aber mit dem Studium der diskontinuierlichen endlichen wie unendlichen Gruppen, die er für eine Fülle der verschiedenartigsten Untersuchungen geometrischer, zahlentheoretischer und funktionentheoretischer Art verwendete. Seine Untersuchungen sind in folgenden Lehrbüchern zusammengefaßt: *Vorlesungen über das Ikosaëder und die Auflösung der Gleichungen vom 5. Grade* (Leipzig 1884). *Vorlesungen über die*

Theorie der elliptischen Modulfunktionen, herausg. von Fricke, 2 Bde., Leipzig 1890/92. Fricke u. Klein, *Vorlesungen über automorphe Funktionen*, Bd. 1, Leipzig 1897, Bd. 2, 1. Lieferung 1901.

Außer den schon zitierten Werken kommen für die in diesem Kapitel zu behandelnden Gegenstände hauptsächlich folgende Lehrbücher in Frage:

Für das Gesamtgebiet: H. Weber, *Lehrbuch der Algebra*. Bd. 1 u. 2, Braunschweig, 2. Aufl. 1898 u. 1899.

Für Permutationsgruppen: E. Netto, *Substitutionentheorie und ihre Anwendungen auf Algebra*, Leipzig 1882.

Bianchi, *Lezioni sulla teoria dei gruppi di sostituzioni*, Pisa 1900.

Für abstrakte Gruppen: J. A. de Séguier, *Éléments de la théorie des groupes abstraits*, Paris 1904.

Für endliche abstrakte Gruppen und Permutationsgruppen: W. Burnside, *Theory of groups of finite order*, Cambridge 1897, E. Netto, *Gruppen- und Substitutionentheorie, Samml. Schubert* **55**, Leipzig 1908.

Eine eingehende Zusammenstellung der Literatur und der Lehrsätze über Permutationsgruppen und endliche abstrakte Gruppen gibt B. S. Easton, *The constructive development of group-theory with a bibliography*, Boston 1902. Sehr schätzenswerte Führer in der neueren gruppentheoretischen Literatur sind die Referate von G. A. Miller, *Bull. Am. M. S.* (2) **5**, 227—249 (1899), (2) **9**, 106—123 (1902), (2) **14**, 78—91, 124—133 (1907), (2) **7**, 121—130 (1900), sowie von L. E. Dickson, ebenda (2) **6**, 13—27 (1899). Schließlich sei noch auf die Artikel *„Endliche diskrete Gruppen“* von H. Burkhardt und *„Endliche Gruppen linearer Substitutionen“* von A. Wiman in der *Encyklopädie der math. Wiss.* **1**, 208 u. 522 verwiesen.

Eine *Gruppe* läßt sich auf folgende Weise *definieren:* Für ein System $\mathfrak{G}$ von Elementen — man sagt auch Dingen, Operatoren — $A, B, C, \ldots$ sei irgendeine Vorschrift gegeben, die aus zwei beliebigen gleichen oder ungleichen Elementen A und B eindeutig stets ein drittes C bestimmt. Man schreibt $C = AB$. Das Verfahren wird als Komposition oder Multiplikation bezeichnet, C heißt das Produkt von A und B oder auch das Resultat der Komposition von A und B.

Die Elemente von $\mathfrak{G}$ bilden eine Gruppe, wenn sie, außer

daß sie untereinander komponiert werden können, noch die folgenden vier Postulate erfüllen:

1. Das Produkt irgend zweier Elemente von $\mathfrak{G}$ gehört $\mathfrak{G}$ an.

2. Die Produktbildung ist assoziativ, d. h. sind A, B, C irgend drei Elemente aus $\mathfrak{G}$, so soll $A(BC) = (AB)C$ sein.

3. In $\mathfrak{G}$ gibt es wenigstens ein Element J, so daß für jedes Element A von $\mathfrak{G}$ die Gleichung $AJ = A$ gilt. (J heißt rechtshändiges Einheitselement.)

4. Existieren Elemente J, so soll für ein besonderes J und für jedes A die Gleichung $AX = J$ durch ein Element von $\mathfrak{G}$ lösbar sein.

Die angegebenen vier Postulate sind *voneinander unabhängig*, d. h. es läßt sich aus drei von ihnen niemals das vierte als beweisbarer Lehrsatz ableiten. Die vier Postulate bleiben auch voneinander unabhängig, wenn man zu ihnen noch ein fünftes über die Anzahl der verschiedenen Elemente von $\mathfrak{G}$ hinzufügt: Die Anzahl der verschiedenen Elemente von $\mathfrak{G}$

5_1) sei gleich n, wobei n eine beliebige endliche Zahl ist, oder

5_2) sie bestehe aus einer abzählbar unendlichen Menge oder

5_3) sie bilde eine nicht abzählbare unendliche Menge.

Die gegebene Definition einer abstrakten Gruppe durch voneinander unabhängige Postulate ist die Moore-Dicksonsche (E. H. Moore, *Trans. Am. M. S.* **3**, 485 (1902), **5**, 549 (1904), **6**, 179, L. E. Dickson, ebenda **6**, 198 (1905)). Wegen anders gefaßter Definitionen einer abstrakten Gruppe vgl. Huntington, *Trans. Am. M. S.* **6**, 181 (1905).

Man beweist: *Eine auf die obige Weise definierte Gruppe $\mathfrak{G}$ enthält nur ein einziges Einheitselement;* es läßt sowohl als linkshändiger oder vorderer Faktor wie als rechtshändiger oder hinterer Faktor jedes Element von $\mathfrak{G}$ ungeändert. *Dieses einzige Einheitselement wird im folgenden ausnahmslos mit E oder 1 bezeichnet.*

Auf Grund der eingeführten Postulate beweist man ferner: Zu jedem Element A von $\mathfrak{G}$ gibt es in $\mathfrak{G}$ ein und auch nur ein Element X, das gleichzeitig den zwei Gleichungen $AX = E$ und $XA = E$ genügt. Dieses Element X wird mit A^{-1} bezeichnet und heißt das zu A *inverse* oder *reziproke* Element. Es ist also $AA^{-1} = A^{-1}A = E$.

Für das reziproke Element eines Produktes AB gilt die Relation

$$(AB)^{-1} = B^{-1}A^{-1}.$$

Setzt man ein Element A α-mal als Faktor, so wird das Resultat der Komposition mit A^α bezeichnet und die α^{te} *Potenz von* A genannt. Unter A^0 soll das Einheitselement E verstanden werden. Setzt man A^{-1} α-mal als Faktor, so entsteht ein Element, das mit $A^{-\alpha}$ bezeichnet wird. $(A^{-1})^\alpha = A^{-\alpha}$ ist das reziproke Element von A^α. Für positive und negative ganzzahlige Werte von α und β gilt ebenso wie für verschwindende α und β die Relation

$$A^\alpha A^\beta = A^\beta A^\alpha = A^{\alpha+\beta}.$$

Eine Gruppe $\mathfrak{G}$ mit einer endlichen Anzahl verschiedener Elemente heißt eine *endliche Gruppe*. Besitzt die endliche Gruppe $\mathfrak{G}$ n verschiedene Elemente, so heißt n die *Ordnung der Gruppe* (Cauchy bedient sich schon der Bezeichnung „ordre", H. Weber, *Algebra* **2**, 4 verwendet abweichend vom üblichen Sprachgebrauch Grad für Ordnung), und man *schreibt* $\mathfrak{G}_n$. Eine Gruppe mit unendlich vielen Elementen heißt eine *unendliche Gruppe*.

Bei einer Gruppe ist das Produkt AB zweier Elemente im allgemeinen von dem Produkt BA verschieden. Ist das Produkt zweier Gruppenelemente A und B unabhängig von der Reihenfolge der Faktoren, ist also $AB = BA$, so heißen die Elemente A und B *vertauschbar* oder *kommutativ*. Jedes Gruppenelement ist mit seinen positiven wie negativen Potenzen vertauschbar. Das Einheitselement ist mit allen Gruppenelementen vertauschbar.

Eine Gruppe, die nur aus vertauschbaren Elementen besteht, heißt eine *vertauschbare, kommutative* oder *Abelsche* Gruppe. Der letztere Name ist mit Rücksicht auf die Eigenschaften der von Abel untersuchten besonderen Klasse algebraisch auflösbarer Gleichungen gegeben worden, die eine vertauschbare Galoissche Gruppe haben (Abel, *Mémoire sur une classe particulière d'équations résolubles algébriquement, Journ. f. Math.* **4**, 131 (1829), *Œuvres par Sylow et Lie* **1**, 478, deutsche Ausgabe mit Anm. von A. Loewy in *Ostwalds Klassikern der exakten Wiss.* No. 111).

Eine Gruppe, die nur aus den Potenzen eines einzigen Elementes besteht, also bloß die Elemente:

$$\ldots, A^{-2}, A^{-1}, A^0, A, A^2, \ldots$$

umfaßt, heißt *zyklisch*. Es gibt sowohl endliche als auch un-

endliche zyklische Gruppen. Jede zyklische Gruppe ist eine kommutative Gruppe.

Ein Beispiel für eine nicht kommutative unendliche Gruppe gibt die Gesamtheit aller Matrices gleichen Grades von nichtverschwindenden Determinanten, wenn man sie komponiert; Einheitselement der Gruppe ist hierbei die Einheitsmatrix (vgl. Kap. II, § 6). Weitere Beispiele nichtkommutativer unendlicher Gruppen sind: die Gesamtheit aller Matrices gleichen Grades, deren Determinanten den absoluten Betrag 1 haben, die Gesamtheit aller Matrices gleichen Grades, deren Determinanten den Wert $+1$ haben, die Gesamtheit aller Matrices gleichen Grades, die orthogonale Substitutionen bestimmen, die Gesamtheit aller Matrices, deren Koeffizienten rationale Zahlen oder algebraische Zahlen sind, falls man sie nach dem für Matrices geltenden Kalkül komponiert. Einheitselement bei diesen Gruppen ist stets die Einheitsmatrix.

Durch jedes System höherer komplexer Zahlen werden Gruppen bestimmt (vgl. S. 99ff.).

Als geometrisches Beispiel für eine unendliche nichtkommutative Gruppe kann die schon oben erwähnte Gesamtheit aller kollinearen Umformungen des Raumes dienen. Jede Raumtransformation, die Punkte in Punkte überführt, ist eine eindeutig umkehrbare Operation. Werden zwei solche Transformationen nacheinander ausgeführt, so läßt sich ihr Resultat stets durch eine dritte derselben Art erzielen. Für die Zusammensetzung von Kollineationen gilt offenbar auch das assoziative Gesetz. Die Gesamtheit der kollinearen Umformungen des Raumes hat also Gruppencharakter; man bezeichnet diese Gruppe auch als *„allgemeine projektive Gruppe"*. Die allgemeine Form einer einzelnen projektiven Transformation lautet für unseren R_3 in unhomogenen Punktkoordinaten x, y, z:

$$x = \frac{a_{11}x' + a_{12}y' + a_{13}z' + a_{14}}{a_{41}x' + a_{42}y' + a_{43}z' + a_{44}},$$

$$y = \frac{a_{21}x' + a_{22}y' + a_{23}z' + a_{24}}{a_{41}x' + a_{42}y' + a_{43}z' + a_{44}},$$

$$z = \frac{a_{31}x' + a_{32}y' + a_{33}z' + a_{34}}{a_{41}x' + a_{42}y' + a_{43}z' + a_{44}},$$

wobei die Konstanten a_{ik} nur der Bedingung zu genügen haben, daß die Determinante $|a_{ik}|$ $(i, k = 1, 2, 3, 4)$ nicht verschwinden darf.

Die allgemeine projektive Gruppe umfaßt die Hauptgruppe wie die Bewegungsgruppe (vgl. oben S. 171).

Beispiele für endliche, nichtkommutative Gruppen werden im folgenden in der Form von Permutationsgruppen, Kongruenzgruppen und endlichen Gruppen linearer homogener Substitutionen behandelt werden. Wir begnügen uns damit, an dieser Stelle auf ein durch die Geometrie geliefertes Beispiel hinzuweisen, nämlich auf gewisse in der Bewegungsgruppe unseres Raumes enthaltene Gruppen, die endlichen Gruppen der regulären Körper. Die Gesamtheit der Drehungen, die einen regulären Körper mit sich selbst zur Deckung bringen, bildet eine Gruppe. Das Tetraëder wird durch 12, das Oktaëder und der Würfel durch 24, das Ikosaëder und das Dodekaëder durch 60 Drehungen mit sich selbst zur Deckung gebracht.

Eine unendliche kommutative Gruppe bildet beispielsweise die Gesamtheit unserer gemeinen komplexen Zahlen $a + ib$, wenn man die gewöhnliche Addition als Komposition der Elemente ansieht. Einheitselement dieser Gruppe ist die Null. Eine unendliche kommutative Gruppe wird auch bei Ausschluß der Null von der Gesamtheit unserer gemeinen komplexen Zahlen $a + ib$ gebildet, wenn man die gewöhnliche Multiplikation als Komposition der Elemente auffaßt. Einheitselement der Gruppe ist hierbei die Zahl 1.

Die bei dem System der gemeinen komplexen Zahlen entgegentretende doppelte Verknüpfungsfähigkeit der Elemente mit Gruppencharakter führt zum Begriff des *Feldes* oder *Körpers* oder *Rationalitätsbereiches*. Wir betrachten ein System $\mathfrak{C}$ von Elementen, die sich auf zwei Arten verknüpfen lassen. In bezug auf die erste Verknüpfung, die als Addition bezeichnet wird, sollen die Elemente von $\mathfrak{C}$ eine Gruppe bilden. Für die zweite Art der Komposition, die Multiplikation heiße, wird verlangt, daß die Elemente von $\mathfrak{C}$, wenn man die Einheit der additiven Gruppe, die 0 heiße, ausschließt, eine kommutative Gruppe bilden. Die Multiplikation der Elemente von $\mathfrak{C}$ mit 0 soll nach der Relation $0 \cdot x = x \cdot 0 = 0$ statthaben. Schließlich soll die Multiplikation der Elemente von $\mathfrak{C}$ in Verbindung mit der Addition distributiv sein. $a(b + c) = ab + ac$. Genügen die Elemente von $\mathfrak{C}$ den angeführten Bedingungen, so ist, wie man beweisen kann, auch die bei der additiven Verknüpfung der Elemente von $\mathfrak{C}$ entstehende Gruppe kommutativ (vgl. Hilbert, *Math.-Ver.* 8, 183 (1899)). Ein System $\mathfrak{C}$, dessen Elemente den obigen Bedingungen genügen, heißt ein *Feld*. *Körper* oder

Rationalitätsbereich. Kurz kann das Feld als ein in sich abgeschlossenes Elementensystem definiert werden, für das die gewöhnlichen Sätze der Algebra gelten. Abstrakte Definitionen des Feldes geben H. Weber, *Math. Ann.* **43**, 526 (1893), Dickson, *Trans. Am. M. S.* **4**, 13 (1903), **6**, 198 (1905), Huntington, ebenda **4**, 31 (1903) und **6**, 181 (1905). Unendliche Felder sind: die rationalen Zahlen, alle reellen Zahlen, alle algebraischen Zahlen, die Gesamtheit aller Zahlen, wenn man sie in gewöhnlicher Weise additiv und multiplikativ verknüpft. Die abstrakten Definitionen für den Körper lassen sich auch zur Einführung der gemeinen komplexen Zahlen unseres Zahlensystems verwenden. Vgl. Huntington, *A set of postulates for real algebra, for ordinary complex algebra, Trans. Am. M. S.* **6**, 17 und **6**, 209 (1905).

Fundamental ist die Existenz *endlicher Felder* oder *endlicher Körper*, d. h. solcher mit nur endlich vielen verschiedenen Elementen.

Ist p eine positive Primzahl, so erhält man auf folgende Weise ein endliches Feld mit p verschiedenen Elementen: Wir teilen die ganzen positiven Zahlen mod p in p Klassen, so daß alle mod p kongruenten Zahlen derselben Klasse angehören, während mod p inkongruente Zahlen in verschiedene Klassen fallen:

$$\begin{aligned}
K_0 &: \quad 0, \quad p, \quad 2p, \quad 3p, \quad \ldots \\
K_1 &: \quad 1, \quad p+1, \quad 2p+1, \quad 3p+1, \quad \ldots \\
K_2 &: \quad 2, \quad p+2, \quad 2p+2, \quad 3p+2, \quad \ldots \\
&\vdots \qquad\qquad\qquad\qquad\qquad\qquad \vdots \\
K_{p-1} &: p-1, \quad 2p-1, \quad 3p-1, \quad 4p-1, \quad \ldots
\end{aligned}$$

Wir definieren Addition und Multiplikation der Klassen nach dem Gesetz $K_a + K_b = K_s$, $K_a \cdot K_b = K_m$, wobei s und m die kleinsten positiven Reste mod p der Zahlen $a+b$ bezw. $a \cdot b$ bedeuten. Die Klassen $K_0, K_1, \ldots, K_{p-1}$ oder ihre Repräsentanten, die Zahlen $0, 1, 2, \ldots, p-1$, bilden ein endliches Feld von p Elementen. Man bezeichnet es auch als $GF[p]$.

Ein endliches Feld der Ordnung p^n (p Primzahl) läßt sich mittelst Kongruenzen durch ganze Funktionen einer Variablen mit ganzzahligen Koeffizienten einführen. Zwei ganze Funktionen $f(x)$ und $f_1(x)$ einer Variablen x mit ganzzahligen Koeffizienten heißen *kongruent* in bezug auf die Primzahl p,

wenn in der ganzen Funktion $f(x) - f_1(x)$ jeder Koeffizient durch die Primzahl p ohne Rest teilbar ist; alsdann schreibt man $f(x) \equiv f_1(x) \pmod{p}$. Für die hier vorliegenden Zwecke heißt eine ganze Funktion mit ganzzahligen Koeffizienten vom Grad n, wenn x^n die höchste Potenz ist, deren Koeffizient nicht durch p teilbar ist. Die ganze ganzzahlige Funktion $P(x)$ vom n^{ten} Grad, bei der x^n die Einheit als Koeffizient habe, heißt *irreduzibel*, falls keine Kongruenz $P(x) \equiv f_1(x) f_2(x) \pmod{p}$ möglich ist, wobei $f_1(x)$ und $f_2(x)$ ganze ganzzahlige Funktionen niederen als n^{ten} Grades bedeuten.

Jede ganze ganzzahlige Funktion $G(x)$ läßt sich vermittels der Funktion $P(x)$ in die Form:

$$G(x) = P(x)\, Q(x) + R(x)$$

bringen, wobei der Rest $R(x)$ eine ganze Funktion mit ganzzahligen Koeffizienten von niedrigerem als n^{tem} Grad ist. Sind $G_1(x)$ und $G_2(x)$ zwei ganze Funktionen mit ganzzahligen Koeffizienten und ihre durch Division mit $P(x)$ gewonnenen Reste $R_1(x)$ und $R_2(x)$ kongruent mod p, so werden $G_1(x)$ und $G_2(x)$ als *kongruent in bezug auf den Doppelmodul p, P* bezeichnet. Man schreibt

$$G_1(x) \equiv G_2(x) \pmod{p,\ P(x)}.$$

In bezug auf den Doppelmodul p, P zerfallen die ganzen Funktionen mit ganzzahligen Koeffizienten in p^n Klassen inkongruenter Funktionen; jede von diesen wird durch eine der p^n ganzen Funktionen:

$$t_0 + t_1 x + t_2 x^2 + \cdots + t_{n-1} x^{n-1}$$

repräsentiert; hierbei haben $t_0, t_1, t_2, \ldots, t_{n-1}$ die Reihe der Zahlen $0, 1, 2, \ldots, p-1$ *zu durchlaufen.* Addiert und multipliziert man die ganzen Funktionen mit ganzzahligen Koeffizienten, so ist sowohl ihre Summe als auch ihr Produkt einer der p^n angegebenen untereinander inkongruenten Funktionen kongruent in bezug auf den Doppelmodul (p, P). *Die p^n Klassen* mod (p, P) *inkongruenter ganzer Funktionen mit ganzzahligen Koeffizienten bilden ein endliches Feld von p^n Elementen.* Man bezeichnet es (E. H. Moore, *Math. papers of the Chicago Congress*, 1893, publ. 1896, S. 211) Galois zu Ehren, der (*Sur la théorie des nombres* (1830), *Œuvres*, p. 15) zuerst endliche Felder mit p^n Elementen verwendete, als $GF[p^n]$. Statt der

oben benützten p^n Funktionen kann man sich auch der von Galois a. a. O. eingeführten sogenannten *Galoisschen Imaginären* oder (H. Weber, *Algebra* 2, 305) der Wurzeln der Gleichung $P = 0$ bedienen. Das $GF[p^n]$ ist Gegenstand des Werkes von L. E. Dickson, *Linear groups with an exposition of the Galois field theory,* Leipzig 1901, vgl. ferner J. A. Serret, *Algèbre supérieure,* Paris 1866, 2, 121, de Séguier, *Éléments de la théorie des groupes abstraits,* Paris 1904, p. 28.

Ein endliches Feld von p^n Elementen liefert auch die Idealtheorie mittels Kongruenzen nach Primidealen n^{ten} Grades. Ist $\mathfrak{p}$ ein Primideal n^{ten} Grades im Körper $\mathfrak{R}$, also die Norm von $\mathfrak{p}$ gleich der n^{ten} Potenz der rationalen Primzahl p, so ist die Anzahl der nach dem Ideal $\mathfrak{p}$ untereinander inkongruenten ganzen Zahlen des Körpers $\mathfrak{R}$ gleich p^n. Diese p^n Zahlen bilden ein Feld mit p^n Elementen. Man kann auch stets einen Körper $\mathfrak{R}$ mit Primidealen $\mathfrak{p}$ finden, deren Norm einen vorgegebenen Wert p^n hat; im Kreiskörper der $p^n - 1^{\text{ten}}$ Einheitswurzeln ist nämlich die Primzahl p in Primideale n^{ten} Grades zerlegbar (vgl. Hilbert, *Die Theorie der algebraischen Zahlkörper, Math.-Ver.* (1897), Satz 125, S. 333).

Für die endlichen Felder gilt folgender Fundamentalsatz von E. H. Moore (*Math. papers of the Chicago Congress,* 1893, S. 220, L. E. Dickson, *Linear groups,* S. 13, E. H. Moore, *The decennial publications of the university of Chicago,* Vol. IX, S. 7, Chicago 1903): *Die Anzahl der verschiedenen Elemente irgendeines endlichen Feldes kann nur gleich einer Primzahl oder gleich einer ihrer Potenzen sein. Jedes mögliche endliche Feld hat in einem $GF[p^n]$ seinen Vertreter. Das $GF[p^n]$ ist unabhängig von der speziellen Wahl der irreduziblen Funktion $P(x)$; es wird durch p und n völlig charakterisiert* (letzteres Resultat eigentlich schon bei Galois, *Œuvres,* p. 17). Jedes endliche Feld $\mathfrak{C}$ mit p^n Elementen kann also mit jedem endlichen Feld $\mathfrak{C}'$ mit p^n Elementen derartig in eine eindeutige Beziehung gesetzt werden, daß, wenn a', b', c' Elemente aus $\mathfrak{C}'$ sind, die den Elementen a, b, c aus $\mathfrak{C}$ entsprechen, sich aus der Relation $a + b = c$ die Gleichung $a' + b' = c'$ und umgekehrt ergibt und aus $ab = c$ die Beziehung $a'b' = c'$ und umgekehrt folgt.

§ 2. Allgemeines über abstrakte Gruppen.

Eine Gruppe $\mathfrak{H}$ heißt eine *Untergruppe* (Lie, Netto, Frobenius) ein *Divisor* (Galois, *Œuvres*, p. 58) oder ein *echter Teiler* (Weber, *Algebra* 2, 7) einer Gruppe $\mathfrak{G}$, wenn sämtliche Elemente von $\mathfrak{H}$ auch der Gruppe $\mathfrak{G}$ angehören und $\mathfrak{G}$ außer den Elementen von $\mathfrak{H}$ wenigstens noch ein Element enthält.

Das Einheitselement ist eine Untergruppe jeder Gruppe. Ist A irgendein Element einer Gruppe $\mathfrak{G}$, so bilden die Elemente

$$\ldots A^{-2}, A^{-1}, A^0, A, A^2, \ldots$$

eine *zyklische Untergruppe von* $\mathfrak{G}$. Hat diese Reihe lauter verschiedene Elemente, so heißt A ein *Element unendlich hoher Ordnung.* Hat die obige Reihe nicht lauter untereinander verschiedene Elemente, so wird eine Potenz des Elementes A gleich dem Einheitselement. In diesem Falle heißt *,das Element A von endlicher Ordnung.* Ist a die kleinste positive Zahl, für die $A^a = 1$ wird, so heißt a die *Ordnung* oder die *Periode des Elementes* A. (H. Weber, *Algebra* 2, 11 sagt Grad des Elementes A.)

Ist A ein Element der endlichen Ordnung a, so enthält die Reihe: $A^0, A^1, A^2, A^3, \ldots, A^{a-1}$ a voneinander verschiedene Elemente und erschöpft alle positiven wie negativen Potenzen von A. (Diese Sätze finden sich bereits bei Abel, vgl. *Ostwalds Klassiker d. exakten Wiss.* Nr. 111, S. 6.)

Alle Elemente, die zwei Gruppen $\mathfrak{G}_1$ und $\mathfrak{G}_2$ gemeinschaftlich angehören, bilden wiederum eine Gruppe; sie heißt der *größte gemeinschaftliche Teiler* oder nach Study (H. Weber, *Algebra* 2, 10) der *Durchschnitt von* $\mathfrak{G}_1$ *und* $\mathfrak{G}_2$. Zwei Gruppen, deren Durchschnitt nur aus dem Einheitselement besteht, heißen *teilerfremd.*

Die Gleichung $\mathfrak{A} = A + B + C + \cdots$ soll ausdrücken, die Elemente $A, B, C, \ldots$ sind zu einem System $\mathfrak{A}$ zusammengefaßt; die Gleichung

$$\mathfrak{R} = \mathfrak{A} + \mathfrak{B} + \mathfrak{C} + \cdots$$

bedeutet, die Systeme $\mathfrak{A}, \mathfrak{B}, \mathfrak{C}, \ldots$ sollen zu einem weiteren System $\mathfrak{R}$ vereinigt werden. Durchläuft A alle Elemente des Systems $\mathfrak{A}$ und B alle Elemente des Systems $\mathfrak{B}$, so durchläuft AB ein System von Elementen, das mit $\mathfrak{AB}$ bezeichnet sei.

Die Bildung $\mathfrak{AB}$ bezeichnet man als *Komposition von Systemen.* (H. Weber, *Algebra* **2**, 13 spricht von der *Komposition der Teile.*) Besteht $\mathfrak{B}$ nur aus einem einzigen Element B, so wird $\mathfrak{A}B$ für $\mathfrak{AB}$ geschrieben; ebenso verwendet man, wenn $\mathfrak{A}$ nur aus einem einzigen Element A besteht, die Bezeichnung $A\mathfrak{B}$. Für *Gruppen* und *Systeme* werden im folgenden *große deutsche,* für *einzelne Elemente große lateinische Buchstaben* verwendet. (Vgl. hierzu Frobenius, *Sitzungsb. d. Berl. Akad.* (1895), 163.) Ein *Gleichheitszeichen* $\mathfrak{A} = \mathfrak{B}$ soll besagen, daß jedes Element des einen Systems auch in dem anderen vorkommt; hierbei soll jedoch nicht verlangt sein, daß in $\mathfrak{A}$ oder $\mathfrak{B}$ wiederholt auftretende Elemente sich in dem einen System ebenso häufig als in dem anderen vorfinden. $\mathfrak{A} > \mathfrak{B}$ besagt: Das System $\mathfrak{A}$ enthält alle Elemente von $\mathfrak{B}$, aber außerdem auch noch solche, die nicht in $\mathfrak{B}$ auftreten.

Zwei Systeme $\mathfrak{A}$ *und* $\mathfrak{B}$ *heißen vertauschbar, wenn* $\mathfrak{AB} = \mathfrak{BA}$ *ist. d. h. jedes Element des Systemes* $\mathfrak{AB}$ *in* $\mathfrak{BA}$ *auftritt und umgekehrt.*

Ein Element B *heißt mit einem System* $\mathfrak{A}$ *vertauschbar, wenn die Gleichung* $\mathfrak{A}B = B\mathfrak{A}$ *statthat.*

Jedes in einer Gruppe $\mathfrak{G}$ enthaltene Element G ist mit der Gruppe $\mathfrak{G}$ vertauschbar: $G\mathfrak{G} = \mathfrak{G}G$.

Sind $\mathfrak{G}_1$ und $\mathfrak{G}_2$ zwei Gruppen, so heißt die kleinste Gruppe $\mathfrak{C}$, die sowohl alle Elemente von $\mathfrak{G}_1$ als auch von $\mathfrak{G}_2$ umfaßt, *das kleinste gemeinsame Vielfache von* $\mathfrak{G}_1$ *und* $\mathfrak{G}_2$. Es ist $\mathfrak{C} \geq \mathfrak{G}_1\mathfrak{G}_2$; das Gleichheitszeichen gilt dann und nur dann, wenn $\mathfrak{G}_1$ und $\mathfrak{G}_2$ vertauschbare Gruppen sind, also $\mathfrak{G}_1\mathfrak{G}_2 = \mathfrak{G}_2\mathfrak{G}_1$ ist. Eine Gruppe $\mathfrak{C}$, die zwei derartige Untergruppen $\mathfrak{G}_1$ und $\mathfrak{G}_2$ enthält, daß $\mathfrak{C} = \mathfrak{G}_1\mathfrak{G}_2 = \mathfrak{G}_2\mathfrak{G}_1$ ist, heißt *zerlegbar* (décomposable). Vgl. Frobenius, *Sitzungsb. d. Berl. Akad.* (1895), 166; E. Maillet, *Bull. soc. math.* **28**, 7 (1900).

Sind $\mathfrak{G}_1$ und $\mathfrak{G}_2$ zwei teilerfremde Gruppen und ist jedes Element von $\mathfrak{G}_1$ mit jedem Element von $\mathfrak{G}_2$ vertauschbar, so heißt nach Hölder (*Math. Ann.* **34**, 36 (1889), ebenda **43**, 330 (1893)) das kleinste gemeinsame Vielfache $\mathfrak{C}$ von $\mathfrak{G}_1$ und $\mathfrak{G}_2$ das *direkte Produkt von* $\mathfrak{G}_1$ *und* $\mathfrak{G}_2$; die Gruppen $\mathfrak{G}_1$ und $\mathfrak{G}_2$ sind invariante Untergruppen von $\mathfrak{C}$. Eine Gruppe, die als direktes Produkt zweier Gruppen darstellbar ist, heißt *zerfallend.* (Dyck, *Math. Ann.* **17**, 482 (1880); Hölder, *Math. Ann.* **34**, 36 (1889), ebenda **43**, 335 (1893).) Zerfallende Gruppen sind eine spezielle Art zerlegbarer Gruppen.

$\mathfrak{H}_1$ und $\mathfrak{H}_2$ seien zwei gleiche oder verschiedene Gruppen, R irgendein Element. Betrachtet man das System $\mathfrak{H}_1 R \mathfrak{H}_2$ und ersetzt R durch irgendein anderes Element R', das dem System $\mathfrak{H}_1 R \mathfrak{H}_2$ angehört, so enthalten die zwei Systeme $\mathfrak{H}_1 R \mathfrak{H}_2$ und $\mathfrak{H}_1 R' \mathfrak{H}_2$ die nämlichen Elemente. Hieraus folgt: Haben zwei Systeme $\mathfrak{H}_1 R \mathfrak{H}_2$ und $\mathfrak{H}_1 S \mathfrak{H}_2$ ein Element gemeinsam, so stimmen sie in allen Elementen überein. Tritt das letztere ein, so heißen *die Elemente R und S äquivalent nach dem Doppelmodul* $(\mathfrak{H}_1, \mathfrak{H}_2)$. Die Äquivalenz nach einem Doppelmodul ist eingeführt von Frobenius, *Journ. f. Math.* **101**, 273 (1887); vgl. ferner Dedekind, *Gött. Nachr.* (1894), 275, Frobenius, *Sitzungsb. d. Berl. Akad.* (1895), 167, H. Weber, *Algebra* **2**, 22, de Séguier, *Groupes abstraits*, S. 61.

Ist $\mathfrak{G}$ eine Gruppe mit den Untergruppen $\mathfrak{H}_1$ und $\mathfrak{H}_2$, so kann man $\mathfrak{G}$ in Systeme zerlegen:

$$(1) \qquad \mathfrak{G} = \mathfrak{H}_1 G_1 \mathfrak{H}_2 + \mathfrak{H}_1 G_2 \mathfrak{H}_2 + \mathfrak{H}_1 G_3 \mathfrak{H}_2 + \cdots;$$

hierbei haben niemals zwei auf der rechten Seite der Gleichung (1) stehende Systeme $\mathfrak{H}_1 G_i \mathfrak{H}_2$ und $\mathfrak{H}_1 G_k \mathfrak{H}_2$ ein Element gemeinsam. Die Elemente $G_1 = 1, G_2, G_3, \ldots$ heißen ein *vollständiges System nicht äquivalenter Elemente* oder ein *vollständiges Restsystem der Gruppe $\mathfrak{G}$ nach dem Doppelmodul* $(\mathfrak{H}_1, \mathfrak{H}_2)$; hierbei kann jedes Element G_i durch jedes andere Element von $\mathfrak{H}_1 G_i \mathfrak{H}_2$ ersetzt werden. Ist $\mathfrak{G}$ eine Gruppe mit einer nicht abzählbaren Menge von Elementen, so soll, wie bemerkt sei, durch die Numerierung und verwendete Schreibweise nicht etwa ausgedrückt werden, daß die Systeme $\mathfrak{H}_1 G_1 \mathfrak{H}_2$, $\mathfrak{H}_1 G_2 \mathfrak{H}_2$, $\mathfrak{H}_1 G_3 \mathfrak{H}_2$, ... eine im Sinn von G. Cantor abzählbare Menge sind.[1])

Gleichzeitig mit der Zerlegung (1) besteht die Zerlegung:

$$(2) \qquad \mathfrak{G} = \mathfrak{H}_2 G_1^{-1} \mathfrak{H}_1 + \mathfrak{H}_2 G_2^{-1} \mathfrak{H}_1 + \mathfrak{H}_2 G_3^{-1} \mathfrak{H}_1 + \cdots,$$

bei der die Systeme $\mathfrak{H}_2 G_i^{-1} \mathfrak{H}_1$ und $\mathfrak{H}_2 G_k^{-1} \mathfrak{H}_1$ niemals ein Element gemeinsam haben.

Wählt man bei der Zerlegung einer Gruppe nach einem aus zwei ihrer Untergruppen $\mathfrak{H}_1$ und $\mathfrak{H}_2$ gebildeten Doppel-

1) Diese Festsetzung, daß bei unendlich vielen Elementen im folgenden Numerierung und Punkte keine Abzählbarkeit im Sinne von G. Cantor bedeuten sollen, soll für unsere Betrachtungen allgemein gelten.

modul für die Gruppe $\mathfrak{H}_2$ das Einheitselement, das ja für sich allein eine Gruppe bildet, so geht die Zerlegung (1) in die bereits von *Galois* (*Œuvres*, p. 26) studierte über:

$$(1') \qquad \mathfrak{G} = \mathfrak{H}_1 G_1 + \mathfrak{H}_1 G_2 + \mathfrak{H}_1 G_3 + \cdots;$$

hierbei haben niemals zwei Systeme $\mathfrak{H}_1 G_i$ und $\mathfrak{H}_1 G_k$ ein Element gemeinsam. Die Elemente $G_1 = 1, G_2, G_3, \ldots$ heißen *ein vollständiges System nicht äquivalenter Elemente* oder ein *vollständiges Restsystem der Gruppe* $\mathfrak{G}$ *nach dem Modul* $\mathfrak{H}_1$. In dem Restsystem kann jedes Element G_i $(i = 1, 2, 3, \ldots)$ durch jedes andere Element des Systemes $\mathfrak{H}_1 G_i$ ersetzt werden. Trotzdem nur $\mathfrak{H}_1 G_1 = \mathfrak{H}_1$ eine Gruppe ist, bezeichnet man die Systeme $\mathfrak{H}_1 G_i$ $(i = 1, 2, 3, \ldots)$ nach H. Weber, *Algebra* **2**, 8 als ein *System von Nebengruppen* und spricht von *einer Zerlegung von* $\mathfrak{G}$ *in ein System von Nebengruppen.*

Gleichzeitig mit der Zerlegung (1') besteht die Zerlegung:

$$(2') \qquad \mathfrak{G} = G_1^{-1}\mathfrak{H}_1 + G_2^{-1}\mathfrak{H}_1 + G_3^{-1}\mathfrak{H}_1 + \cdots,$$

bei der zwei Systeme $G_i^{-1}\mathfrak{H}_1$ und $G_k^{-1}\mathfrak{H}_1$ niemals ein Element gemeinsam haben.

Ist bei der Zerlegung (1') die Anzahl der Nebengruppen endlich, besteht also die Zerlegung:

$$(1'') \qquad \mathfrak{G} = \mathfrak{H}_1 G_1 + \mathfrak{H}_1 G_2 + \mathfrak{H}_1 G_3 + \cdots + \mathfrak{H}_1 G_j$$

und infolgedessen auch gleichzeitig:

$$(2'') \qquad \mathfrak{G} = G_1^{-1}\mathfrak{H}_1 + G_2^{-1}\mathfrak{H}_1 + G_3^{-1}\mathfrak{H}_1 + \cdots + G^{-1}\mathfrak{H}_1,$$

so heißt $\mathfrak{H}_1$ *eine Untergruppe von* $\mathfrak{G}$ *von endlichem Index* j *und* j *der Index der Untergruppe.* Existiert keine solche endliche Zahl j, so ist $\mathfrak{H}_1$ eine Untergruppe von *unendlichem Index.* (Poincaré, *Journ. de math.* (4) **3**, 409 (1887).)

Nach Poincaré (a. a. O.) heißen *zwei Gruppen meßbar*, wenn ihr Durchschnitt für jede von ihnen eine Untergruppe von endlichem Index ist.

Sind zwei Gruppen $\mathfrak{G}_1$ und $\mathfrak{G}_2$ mit einer und derselben dritten Gruppe $\mathfrak{G}$ meßbar, so sind $\mathfrak{G}_1$ und $\mathfrak{G}_2$ untereinander meßbar und auch der größte gemeinsame Teiler von $\mathfrak{G}$, $\mathfrak{G}_1$ und $\mathfrak{G}_2$ ist für jede der drei Gruppen eine Untergruppe von endlichem Index.

Ist $\mathfrak{F}$ eine Gruppe, $\mathfrak{G}$ eine Untergruppe von $\mathfrak{F}$ und $\mathfrak{H}$ eine Untergruppe von $\mathfrak{G}$, so gilt die Relation:

$$(\mathfrak{F}, \mathfrak{H}) = (\mathfrak{F}, \mathfrak{G}) \cdot (\mathfrak{G}, \mathfrak{H}),$$

wenn $(\mathfrak{F}, \mathfrak{H})$ (vgl. H. Weber, *Algebra* 2, 9) der (endliche oder unendliche) Index der Untergruppe $\mathfrak{H}$ von $\mathfrak{F}$, und analog $(\mathfrak{F}, \mathfrak{G})$ bzw. $(\mathfrak{G}, \mathfrak{H})$ die Indices der Untergruppen $\mathfrak{G}$ von $\mathfrak{F}$ bzw. $\mathfrak{H}$ von $\mathfrak{G}$ bedeuten.

Ist $\mathfrak{G}_1$ eine beliebige Gruppe und R irgendein Element, das nicht $\mathfrak{G}_1$ anzugehören braucht, so bildet auch das System von Elementen $\mathfrak{G}_2 = R\mathfrak{G}_1 R^{-1}$ eine Gruppe. Man sagt: $\mathfrak{G}_2$ ist eine *transformierte Gruppe von* $\mathfrak{G}_1$ oder *durch Transformation aus* $\mathfrak{G}_1$ hervorgegangen. $\mathfrak{G}_2$ heißt mit $\mathfrak{G}_1$ *ähnlich* (dies ist die älteste an Cauchy anknüpfende Bezeichnung, so bei J. A. Serret, *Algèbre supérieure* 2, 255 (Paris, 3ième éd., 1866), Netto, *Substitutionentheorie*, S. 47), *konjugiert, gleichberechtigt* (Klein, Math. Ann. 14, 430 (1879)) oder *äquivalent*. Die gegebene Definition ist in $\mathfrak{G}_1$ und $\mathfrak{G}_2$ symmetrisch; denn aus $\mathfrak{G}_2 = R\mathfrak{G}_1 R^{-1}$ folgt $\mathfrak{G}_1 = R^{-1}\mathfrak{G}_2 R = R^{-1}\mathfrak{G}_2 (R^{-1})^{-1}$. Zwei Gruppen, die mit einer dritten ähnlich sind, sind es untereinander.

Zwei Untergruppen $\mathfrak{H}_1$ und $\mathfrak{H}_2$ einer Gruppe $\mathfrak{G}$ heißen *in bezug auf* $\mathfrak{G}$ *ähnliche, konjugierte, gleichberechtigte* oder *äquivalente Untergruppen* von $\mathfrak{G}$, wenn es in $\mathfrak{G}$ wenigstens ein Element R gibt, daß die Gruppen $\mathfrak{H}_2$ und $R\mathfrak{H}_1 R^{-1}$ in ihren Elementen übereinstimmen. Die gegebene Definition ist in $\mathfrak{H}_1$ und $\mathfrak{H}_2$ symmetrisch gebaut; denn aus $\mathfrak{H}_2 = R\mathfrak{H}_1 R^{-1}$ folgt $\mathfrak{H}_1 = R^{-1}\mathfrak{H}_2 R = R^{-1}\mathfrak{H}_2 (R^{-1})^{-1}$, und wegen des Gruppencharakters von $\mathfrak{G}$ ist R^{-1} als reziprokes Element von R in $\mathfrak{G}$ enthalten.

Statt in bezug auf $\mathfrak{G}$ ähnlicher oder konjugierter Untergruppen spricht man auch kurz von *ähnlichen* oder *konjugierten Untergruppen von* $\mathfrak{G}$. Dabei ist folgendes zu beachten: Ist $\mathfrak{F}$ eine Gruppe, die $\mathfrak{G}$ als Untergruppe enthält, so können zwei Untergruppen $\mathfrak{H}_1$ und $\mathfrak{H}_2$ von $\mathfrak{G}$ miteinander sehr wohl in bezug auf $\mathfrak{F}$ ähnlich sein, ohne daß es in bezug auf $\mathfrak{G}$ der Fall ist; zwei solche Gruppen $\mathfrak{H}_1$ und $\mathfrak{H}_2$ sind als ähnliche oder konjugierte Untergruppen von $\mathfrak{F}$, nicht aber als ähnliche oder konjugierte Untergruppen von $\mathfrak{G}$ zu bezeichnen.

Jede Untergruppe einer Gruppe $\mathfrak{G}$ ist in bezug auf $\mathfrak{G}$ mit sich selbst ähnlich. Zwei ähnliche Untergruppen einer Gruppe $\mathfrak{G}$,

die in bezug auf $\mathfrak{G}$ einer dritten Untergruppe von $\mathfrak{G}$ ähnlich sind, sind es auch untereinander. Transformiert man eine Untergruppe $\mathfrak{H}$ der Gruppe $\mathfrak{G}$ durch ein vollständiges Restsystem der Gruppe $\mathfrak{G}$ nach dem Modul $\mathfrak{H}$, so erhält man alle mit $\mathfrak{H}$ in bezug auf $\mathfrak{G}$ ähnlichen Untergruppen von $\mathfrak{G}$, diese brauchen nicht alle voneinander verschieden zu sein. Ist $\mathfrak{H}$ im besondern eine Untergruppe von $\mathfrak{G}$ von endlichem Index, so ist die Anzahl der verschiedenen in bezug auf $\mathfrak{G}$ ähnlichen Untergruppen jedenfalls endlich.

Auch für die *einzelnen Elemente* einer Gruppe ist der Begriff der Ähnlichkeit zu definieren: *Zwei Elemente* A und B einer Gruppe $\mathfrak{G}$ heißen *in bezug auf* $\mathfrak{G}$ *ähnlich, konjugiert, gleichberechtigt* oder *äquivalent*, falls es in $\mathfrak{G}$ wenigstens ein Element C gibt, daß $A = CBC^{-1}$ wird. Wie aus $B = C^{-1}AC = C^{-1}A(C^{-1})^{-1}$ folgt, ist die gegebene Definition, da $\mathfrak{G}$ infolge seines Gruppencharakters neben jedem Element C auch das reziproke C^{-1} enthält, in A und B symmetrisch. Von dem Produkt CBC^{-1} sagt man: es ist *durch Transformation aus* B *hervorgegangen.*

Jedes Gruppenelement ist mit sich selbst ähnlich. Sind A_1 und A_2 zwei beliebige Elemente einer Gruppe $\mathfrak{G}$, so sind die zwei Produkte A_1A_2 und $A_2A_1 = A_2(A_1A_2)A_2^{-1}$ miteinander in bezug auf $\mathfrak{G}$ ähnlich.

Zwei Gruppenelemente von $\mathfrak{G}$, die in bezug auf $\mathfrak{G}$ mit einem dritten ähnlich sind, sind es auch untereinander. Hieraus folgt: Alle mit einem Element einer Gruppe $\mathfrak{G}$ in bezug auf $\mathfrak{G}$ ähnlichen Elemente bilden eine Klasse ähnlicher Elemente, so daß zwei Elemente derselben Klasse stets untereinander ähnlich sind. Die Elemente einer Gruppe lassen sich also in *Klassen konjugierter Elemente* einteilen. (Frobenius, *Journ. f. Math.* **100**, 181 (1887).)

Sind C_1 und C_2 zwei in $\mathfrak{G}$ enthaltene Elemente, die beide bewirken, daß $C_1AC_1^{-1} = B$ und $C_2AC_2^{-1} = B$ wird, so ist $C_1^{-1}C_2A = AC_1^{-1}C_2$, d. h. das in $\mathfrak{G}$ befindliche Element $C_1^{-1}C_2$ ist mit A vertauschbar. Hieraus folgt: Sind A und B zwei in bezug auf die Gruppe $\mathfrak{G}$ ähnliche Elemente von $\mathfrak{G}$, d. h. zwei Elemente derselben Klasse, so sind alle in $\mathfrak{G}$ enthaltenen Elemente X, die bewirken, daß $XAX^{-1} = B$ wird, von der Form $C_1\mathfrak{B}$, wobei $\mathfrak{B}$ das System aller mit A vertauschbaren Elemente von $\mathfrak{G}$ bedeutet.

Ist A irgendein Element einer Gruppe $\mathfrak{G}$, so bilden alle in $\mathfrak{G}$ enthaltenen, mit A vertauschbaren Elemente eine Gruppe.

Das System $\mathfrak{B}$ ist also eine Gruppe, und man kann die Gruppe $\mathfrak{G}$ mit Galois nach dem Modul $\mathfrak{B}$ zerlegen:

$$\mathfrak{G} = \mathfrak{B}G_1 + \mathfrak{B}G_2 + \mathfrak{B}G_3 + \cdots$$

Hieraus ergibt sich: *Die Klasse der mit einem Element A einer Gruppe $\mathfrak{G}$ in bezug auf dieses ähnlichen Gruppenelemente besteht aus den Elementen: A, $G_2AG_2^{-1}$, $G_3AG_3^{-1}$, ..., die sämtlich untereinander verschieden sind. $G_1 = 1, G_2, G_3, \ldots$ bedeuten ein vollständiges Restsystem der Gruppe $\mathfrak{G}$ nach dem Modul $\mathfrak{B}$; hierbei ist $\mathfrak{B}$ die Gruppe aller in $\mathfrak{G}$ enthaltenen, mit dem Element A vertauschbaren Elemente. Ist also im besonderen $\mathfrak{B}$ in bezug auf $\mathfrak{G}$ von endlichem Index j, so gibt es innerhalb $\mathfrak{G}$ genau j mit A ähnliche Gruppenelemente.*

Unter Umständen kann auch ein Element A einer Gruppe $\mathfrak{G}$ mit allen Elementen von $\mathfrak{G}$ vertauschbar sein; in diesem Fall fällt $\mathfrak{B}$ mit $\mathfrak{G}$ zusammen. Derartige Elemente sind nur mit sich selbst ähnlich, die Zahl j ist gleich 1. Solche Elemente einer Gruppe $\mathfrak{G}$, die mit jedem Element von $\mathfrak{G}$ vertauschbar sind, heißen *invariante, ausgezeichnete* oder *isolierte* (H. Weber, *Algebra* **2**, 133) *Gruppenelemente.* Das Einheitselement ist invariantes Element jeder Gruppe.

Ist eine Untergruppe $\mathfrak{H}$ von $\mathfrak{G}$ mit allen ihren in bezug auf $\mathfrak{G}$ ähnlichen Untergruppen identisch, so heißt $\mathfrak{H}$ eine *invariante* (Lie, *Math. Ann.* **25**, 77 (1885), nach der dortigen Angabe zuerst eingeführt *Arch. for Math. og. Naturw.* **3**, 457 (1878)) oder eine *ausgezeichnete* (Lie, auf den dieser Ausdruck zurückgeführt wird, lehnt ihn *Math. Ann.* **25**, 77 (1885) ab, die englische Bezeichnung „self-conjugate subgroup" ist als Übersetzung von „ausgezeichneter Untergruppe" von Cole, *Am. J. math.* **9**, 51 (1887) eingeführt worden) oder ein *Normalteiler* (H. Weber, *Algebra* **2**, 12) oder ein *eigentlicher Teiler* von $\mathfrak{G}$.

Eine invariante Untergruppe $\mathfrak{H}$ von $\mathfrak{G}$ kann auch dadurch definiert werden, daß alle Elemente von $\mathfrak{G}$ mit $\mathfrak{H}$ vertauschbar sind, also für jedes Element G von $\mathfrak{G}$ die Gleichung $G\mathfrak{H} = \mathfrak{H}G$ stattfinden muß.

Der Begriff der invarianten Untergruppe ist einer der wichtigsten der ganzen Gruppentheorie; man verdankt ihn Galois. Er spricht, wenn die Zerlegungen

$$\mathfrak{G} = \mathfrak{H}G_1 + \mathfrak{H}G_2 + \mathfrak{H}G_3 + \cdots$$

und

$$\mathfrak{G} = G_1\mathfrak{H} + G_2\mathfrak{H} + G_3\mathfrak{H} + \cdots$$

identisch sind, von einer décomposition propre (*Œuvres*, p. 26); $G_1 = 1, G_2, G_3, \ldots$ bedeuten ein vollständiges Restsystem der Gruppe $\mathfrak{G}$ nach dem Modul $\mathfrak{H}$. Anknüpfend an die Galoissche Ausdrucksweise ist die Bezeichnung „eigentlicher Teiler" (Dedekind, *Math. Ann.* 48, 548 (1897)) entstanden.

Das Einheitselement ist eine invariante Untergruppe jeder Gruppe. Der Durchschnitt aller mit einer Untergruppe $\mathfrak{H}$ einer Gruppe $\mathfrak{G}$ in bezug auf $\mathfrak{G}$ ähnlichen Gruppen ist eine invariante Untergruppe von $\mathfrak{G}$.

Ist $\mathfrak{H}$ eine Untergruppe von $\mathfrak{G}$, $\mathfrak{K}$ eine Untergruppe von $\mathfrak{H}$ und $\mathfrak{K}$ invariante Untergruppe von $\mathfrak{G}$, so ist $\mathfrak{K}$ auch invariante Untergruppe von $\mathfrak{H}$.

Sind $\mathfrak{H}_1$ und $\mathfrak{H}_2$ zwei invariante Untergruppen von $\mathfrak{H}$, so ist ihr Durchschnitt $\mathfrak{D}$ eine invariante Untergruppe von $\mathfrak{H}$, $\mathfrak{H}_1$ und $\mathfrak{H}_2$.

Die Gesamtheit der invarianten Elemente einer Gruppe $\mathfrak{G}$ bildet eine invariante Untergruppe von $\mathfrak{G}$; nach de Séguier, *Groupes abstraits*, S. 57 heißt sie die *Zentrale von* $\mathfrak{G}$. Nur bei einer kommutativen Gruppe, bei dieser aber auch stets, fällt die Zentrale einer Gruppe $\mathfrak{G}$ mit der ganzen Gruppe zusammen.

Jede invariante Untergruppe $\mathfrak{J}$ einer Gruppe $\mathfrak{G}$ bestimmt, wenn $\mathfrak{J}$ nicht das Einheitselement ist, eine neue Gruppe. Sie heißt die *Quotientengruppe von* $\mathfrak{G}$ *und* $\mathfrak{J}$ oder *die zu* $\mathfrak{J}$ *in bezug auf* $\mathfrak{G}$ *komplementäre Gruppe. Sie wird mit* $\mathfrak{G}/\mathfrak{J}$ *bezeichnet.* Man sagt: $\mathfrak{G}$ *ist aus den zwei Gruppen* $\mathfrak{G}/\mathfrak{J}$ *und* $\mathfrak{J}$ *zusammengesetzt. Man bezeichnet* $\mathfrak{G}/\mathfrak{J}$ *auch als Faktorgruppe.* Die Quotientengruppe $\mathfrak{G}/\mathfrak{J}$ ist von C. Jordan, *Bull. soc. math.* **1**, 46 (1873) eingeführt und von Hölder (*Math. Ann.* **34**, 33 (1889)) von neuem aufgestellt worden. Die Quotientengruppe $\mathfrak{G}/\mathfrak{J}$ wird auf folgende Weise gefunden: Man zerlege die Gruppe $\mathfrak{G}$ mittels der invarianten Untergruppe $\mathfrak{J}$ in ein System von Nebengruppen:

$$\mathfrak{G} = \mathfrak{J}G_1 + \mathfrak{J}G_2 + \mathfrak{J}G_3 + \cdots,$$

wobei $G_1 = 1, G_2, G_3, \ldots$ ein vollständiges Restsystem der Gruppe $\mathfrak{G}$ nach dem Modul $\mathfrak{J}$ bilden. Man betrachte die

Nebengruppen $\mathfrak{C}_s = \mathfrak{J} G_s$ $(s = 1, 2, 3, \ldots)$. Die Systeme $\mathfrak{C}_s$ bilden, wenn man sie komponiert, eine Gruppe; ihr Einheitselement ist $\mathfrak{C}_1$. Das Produkt $\mathfrak{C}_k \cdot \mathfrak{C}_l$ zweier Nebengruppen enthält genau die gleichen Elemente wie eine bestimmte Nebengruppe $\mathfrak{C}_m$, so daß $\mathfrak{C}_k \cdot \mathfrak{C}_l = \mathfrak{C}_m$ wird; die Produktbildung ist für die Systeme $\mathfrak{C}$ assoziativ, schließlich sind $\mathfrak{C}_s = \mathfrak{J} G_s$ und $\mathfrak{J} G_s^{-1}$ reziproke Elemente der Gruppe $\mathfrak{G}/\mathfrak{J}$.

Enthält eine Gruppe $\mathfrak{G}$ außer dem Einheitselement keine invariante Untergruppe, so heißt die Gruppe $\mathfrak{G}$ *einfach*. Eine nicht einfache Gruppe heißt *zusammengesetzt*. Zu jeder zusammengesetzten Gruppe $\mathfrak{G}$ gibt es wenigstens eine Quotientengruppe $\mathfrak{G}/\mathfrak{J}$. Der fundamentale Begriff der einfachen Gruppe ist von Galois eingeführt worden. Er spricht von einer „groupe indécomposable“ (*Œuvres*, p. 26).

Eine invariante Untergruppe $\mathfrak{J}$ einer Gruppe $\mathfrak{G}$, die in keiner größeren invarianten Untergruppe von $\mathfrak{G}$ enthalten ist, heißt eine *größte invariante* oder eine *invariante Maximal-Untergruppe* oder ein *größter Normalteiler* oder eine *ausgezeichnete Maximal-Untergruppe von* $\mathfrak{G}$.

Eine Gruppe $\mathfrak{J}$ ist dann und nur dann eine größte invariante Untergruppe einer Gruppe $\mathfrak{G}$, wenn die Quotientengruppe $\mathfrak{G}/\mathfrak{J}$ einfach ist.

Sind $\mathfrak{J}_1$ *und* $\mathfrak{J}_2$ *zwei größte invariante Untergruppen von* $\mathfrak{G}$, *so ist der Durchschnitt* $\mathfrak{D}$ *von* $\mathfrak{J}_1$ *und* $\mathfrak{J}_2$ *sowohl größte invariante Untergruppe von* $\mathfrak{J}_1$ *als auch von* $\mathfrak{J}_2$. *Die Faktorgruppen* $\mathfrak{G}/\mathfrak{J}_1$ *und* $\mathfrak{J}_2/\mathfrak{D}$ *sind holoedrisch isomorph* (vgl. unten); *das gleiche trifft für die Faktorgruppen* $\mathfrak{G}/\mathfrak{J}_2$ *und* $\mathfrak{J}_1/\mathfrak{D}$ *zu.*

Die Gruppe $\mathfrak{G}/\mathfrak{D}$ ist das direkte Produkt der zwei Gruppen $\mathfrak{J}_1/\mathfrak{D}$ und $\mathfrak{J}_2/\mathfrak{D}$. (Vgl. hierzu Hölder, *Math. Ann.* **34**, 36 (1889), Frobenius, *Sitzungsb. d. Berl. Akad.* (1895), 169). Der Begriff der größten invarianten Untergruppe einer Gruppe findet sich bei C. Jordan, *Journ. de math.* (2) **14**, 139 (1869), *Traité*, p. 41.

Eine invariante Untergruppe $\mathfrak{J}$ einer Gruppe $\mathfrak{G}$ heißt eine *charakteristische Untergruppe* von $\mathfrak{G}$, wenn $\mathfrak{J}$ nicht nur in $\mathfrak{G}$, sondern auch in jeder möglichen erweiterten Gruppe invariant ist, die $\mathfrak{G}$ als invariante Untergruppe enthält. Der Begriff der charakteristischen Untergruppe stammt von Frobenius, *Sitzungsb. d. Berl. Akad.* (1895), 183.

Zwei Gruppenelemente A und B bestimmen stets eindeutig ein drittes $A^{-1}B^{-1}AB$; dieses heißt der *Kommutator von A und B*. Die Gesamtheit der Kommutatoren einer Gruppe $\mathfrak{G}$ bildet im allgemeinen für sich keine Gruppe. Da aber nicht nur die Kommutatoren, sondern auch alle aus ihnen durch Produktbildung entstehenden Elemente ebenfalls in der Gruppe $\mathfrak{G}$ enthalten sind, so erzeugen die Kommutatoren, wenn man sie auf jede Weise komponiert, eine Gruppe $\mathfrak{K}$, die nur Elemente aus $\mathfrak{G}$ enthält. $\mathfrak{K}$ heißt die *Kommutatorgruppe* oder die *erste derivierte (abgeleitete) Gruppe von* $\mathfrak{G}$. *Die Gruppe* $\mathfrak{K}$ *ist nicht nur eine invariante, sondern sogar eine charakteristische Untergruppe von* $\mathfrak{G}$.

Ist die Kommutatorgruppe $\mathfrak{K}$ einer Gruppe $\mathfrak{G}$ mit $\mathfrak{G}$ identisch, so heißt $\mathfrak{G}$ eine *perfekte Gruppe*. Jede einfache Gruppe ist a fortiori eine perfekte Gruppe.

Die Kommutatorgruppe $\mathfrak{K}$ einer Gruppe $\mathfrak{G}$ enthält mindestens so viel verschiedene Elemente wie irgendeine Klasse konjugierter Elemente, wenn man $\mathfrak{G}$ in Klassen konjugierter Elemente einteilt. *Ist* $\mathfrak{K}$ *die Kommutatorgruppe einer Gruppe* $\mathfrak{G}$, *so ist* $\mathfrak{G}/\mathfrak{K}$ *eine kommutative Gruppe*. Ist $\mathfrak{J}$ irgendeine invariante Untergruppe von $\mathfrak{G}$, so ist $\mathfrak{G}/\mathfrak{J}$ dann und nur dann eine kommutative Gruppe, wenn $\mathfrak{J}$ die Kommutatorgruppe $\mathfrak{K}$ von $\mathfrak{G}$ zur Untergruppe hat.

Eine charakteristische Eigenschaft einer kommutativen Gruppe ist, daß ihre Kommutatorgruppe das Einheitselement ist.

Die Kommutatorgruppe hat ihren Ausgangspunkt in Lies Theorie der kontinuierlichen Gruppen, Lie-Engel, *Theorie der Transformationsgruppen* **3**, 678 und 770. Von dort, a. a. O., S. 679 stammt auch die Bezeichnung „perfekte Gruppe". Als Literatur über die Kommutatorgruppe ist zu nennen: G. A. Miller, *Quart. Journ.* **28**, 266 (1896), *Bull. Am. M. S.* **4**, 135 (1898), *Am. J* **20**, 277 (1898), G. Frobenius, *Sitzungsb. d. Berl. Akad.* (1896), 1348, Dedekind, *Math. Ann.* **48**, 553 (1897).

Zwei Gruppen $\mathfrak{G}$ und $\mathfrak{G}'$ heißen *holoedrisch isomorph, einfach isomorph, einstufig isomorph* oder kurz *isomorph*, wenn sich ihre Elemente $G_1, G_2, \ldots$ und $G_1', G_2', \ldots$ derartig gegenseitig eindeutig zuordnen lassen, daß, wenn zwei Elemente G_i' und G_k' der Gruppe $\mathfrak{G}'$ den Elementen G_i und G_k der Gruppe $\mathfrak{G}$ entsprechen, das Produkt $G_i'G_k'$ stets dem Produkt G_iG_k zugeordnet ist.

Sind $\mathfrak{G}$ und $\mathfrak{G}'$ isomorphe Gruppen, so ist jeder Untergruppe von $\mathfrak{G}$ eine Untergruppe von $\mathfrak{G}'$ isomorph zugeordnet, den Elementen von $\mathfrak{G}$ entsprechen in $\mathfrak{G}'$ Elemente der nämlichen Ordnung, dem Einheitselement von $\mathfrak{G}$ ist das Einheitselement von $\mathfrak{G}'$ zugeordnet.

Zwei zu einer dritten isomorphe Gruppen sind es nach der Definition des Isomorphismus auch stets untereinander. Jede Gruppe kann daher als Repräsentant aller mit ihr isomorphen Gruppen aufgefaßt werden. Eine spezielle Art des Isomorphismus ist die Ähnlichkeit. Zwei ähnliche Gruppen sind stets isomorph.

Man spricht von einem *meroedrischen Isomorphismus*, auch *mehrstufigen Isomorphismus* zwischen zwei Gruppen $\mathfrak{G}$ und $\mathfrak{G}'$, wenn zwischen ihren Elementen folgende Beziehung statthat: Jedem Element aus $\mathfrak{G}$ entspricht ein und auch nur ein Element aus $\mathfrak{G}'$, jedem Element aus $\mathfrak{G}'$ sind hierdurch ein oder auch mehrere Elemente aus $\mathfrak{G}$ zugeordnet; stets, wenn zwei Elemente G_i' und G_k' aus $\mathfrak{G}'$ den Elementen G_i und G_k aus $\mathfrak{G}$ entsprechen, soll das Produkt $G_i' G_k'$ dem Produkt $G_i G_k$ zugeordnet sein.

Diejenigen Elemente von $\mathfrak{G}$, die dem Einheitselement von $\mathfrak{G}'$ entsprechen, bilden eine invariante Untergruppe $\mathfrak{J}$ von $\mathfrak{G}$. Zerlegt man die Gruppe $\mathfrak{G}$ mit Hilfe der invarianten Untergruppe $\mathfrak{J}$:

$$\mathfrak{G} = \mathfrak{J}G_1 + \mathfrak{J}G_2 + \mathfrak{J}G_3 + \cdots,$$

wobei $G_1, G_2, G_3, \ldots$ ein vollständiges Restsystem der Gruppe $\mathfrak{G}$ nach dem Modul $\mathfrak{J}$ bilden, und entsprechen die Elemente $G_1', G_2', G_3', \ldots$ von $\mathfrak{G}'$ den Elementen $G_1, G_2, G_3, \ldots$ von $\mathfrak{G}$, so ist mit ihnen $\mathfrak{G}'$ erschöpft und allen Elementen des Systemes $\mathfrak{J}G_i$ aus $\mathfrak{G}$ entspricht das Element G_i'. *Die Faktorgruppe* $\mathfrak{G}/\mathfrak{J}$ *ist mit* $\mathfrak{G}'$ *holoedrisch isomorph.*

Die Gruppe $\mathfrak{G}$ heißt mehrstufig oder mehrfach (multiply) isomorph zu $\mathfrak{G}'$ (so bei H. Weber, *Algebra* **2**, 18 sowie bei W. Burnside, *Theory of groups*, S. 36, nach Netto, *Substitutionentheorie*, S. 97, wo der Ausdruck „mehrstufiger Isomorphismus" eingeführt ist, wären $\mathfrak{G}'$ und $\mathfrak{G}$ zu vertauschen), $\mathfrak{G}'$ heißt meroedrisch isomorph zu $\mathfrak{G}$. Enthält $\mathfrak{J}$ unendlich viele Elemente, so sagt man: $\mathfrak{G}$ ist mit $\mathfrak{G}'$ ∞-stufig isomorph. Ist $\mathfrak{J}$ eine endliche Gruppe der Ordnung i, so spricht man von einem i-stufigen Isomorphismus. Der holoedrische und meroedrische Isomorphismus sind, und zwar unter dieser Bezeichnung, eingeführt von C. Jordan, *Traité*, p. 56.

Eine noch *weitergehende Definition des Isomorphismus* als die eben gegebene erhält man nach Capelli, *Giorn. di mat.* **16**, 33 (1878), wenn man die Bedingung fallen läßt, daß jedem Element aus $\mathfrak{G}$ nur *ein einziges* Element aus $\mathfrak{G}'$ entsprechen soll. Zwei Gruppen heißen *allgemein isomorph,* wenn für sie eine derartige wechselseitige Beziehung der Elemente definiert ist, daß jedem Element der einen Gruppe eines oder mehrere der andern Gruppe entsprechen, und falls den Elementen G_i und G_k aus $\mathfrak{G}$ die Elemente G_i' und G_k' aus $\mathfrak{G}'$ entsprechen, auch G_iG_k und $G_i'G_k'$ entsprechende Elemente sind. Sind $\mathfrak{J}_1$, bezw. $\mathfrak{J}_1'$ die Systeme von Elementen aus $\mathfrak{G}$ bezw. $\mathfrak{G}'$, die dem Einheitselement aus $\mathfrak{G}'$ bezw. aus $\mathfrak{G}$ entsprechen, so sind $\mathfrak{J}_1$ und $\mathfrak{J}_1'$ nicht mehr notwendig Gruppen, sondern *Halbgruppen* (Dickson, *Trans. Am. M. S.* **6**, 207 (1905)). Ist eine der Halbgruppen $\mathfrak{J}_1$ oder $\mathfrak{J}_1'$ eine Gruppe, so ist es auch die andere. Sind $\mathfrak{J}_1$ und $\mathfrak{J}_1'$ Gruppen, so sind sie invariante Untergruppen von $\mathfrak{G}$ bezw. $\mathfrak{G}'$ und die Quotientengruppen $\mathfrak{G}/\mathfrak{J}_1$ und $\mathfrak{G}'/\mathfrak{J}_1'$ sind holoedrisch isomorph. (Vgl. Frobenius, *Sitzungsb. d. Berl. Akad.* (1895), 169, de Séguier, *Groupes abstraits,* S. 66.)

Irgendein System $\mathfrak{S}$ von Elementen $A, B, C, \ldots$, die untereinander komponiert werden können, bildet eine *Halbgruppe* (de Séguier, *Groupes abstraits,* S. 8, Dickson, *Trans. Am. M. S.* **6**, 205 (1905)), wenn es die folgenden vier voneinander unabhängigen Postulate erfüllt:

1. Das Produkt irgend zweier Elemente von $\mathfrak{S}$ gehört $\mathfrak{S}$ an.
2. Die Produktbildung ist assoziativ.
3. und 4. Sind G, X und Y irgend drei Elemente aus $\mathfrak{S}$ und ist $GX = GY$ oder $XG = YG$, so soll in jedem der beiden Fälle die Gleichheit $X = Y$ stattfinden.

Eine Halbgruppe mit einer endlichen Anzahl von Elementen ist eine Gruppe. (Vgl. die Definition einer endlichen Gruppe bei H. Weber, *Math. Ann.* **20**, 302 (1882), *Algebra* **2**, 3, Frobenius, *Journ. f. Math.* **100**, 179 (1887).) Eine Halbgruppe braucht nicht das Einheitselement und nicht zu jedem Element das reziproke zu enthalten. Für eine Halbgruppe $\mathfrak{S}$ gilt die symbolische Gleichung $\mathfrak{S}^2 \leqq \mathfrak{S}$, während für eine Gruppe $\mathfrak{S}^2 = \mathfrak{S}$ ist; es gibt auch Halbgruppen, für die $\mathfrak{S}^2 = \mathfrak{S}$ ist.

Für jede Art des Isomorphismus, der hierbei seiner Natur nach unbestimmt gelassen wird, bürgert sich nach F. Klein, *Math. Ann.* **41**, 22 (1893) mehr und mehr die Bezeichnung *Homomorphismus* ein; man beschränkt das Wort „Isomorphismus" dann auf den holoedrischen Isomorphismus. Bezeichnet man im Fall

des meroedrischen Isomorphismus die Gruppen wie früher mit $\mathfrak{G}$ und $\mathfrak{G}'$, so sagt man: $\mathfrak{G}$ ist zu $\mathfrak{G}'$ mehrfach homomorph, $\mathfrak{G}'$ ist mit $\mathfrak{G}$ meroedrisch homomorph. (Vgl. hierzu Burkhardt, *Enzykl. d. math. Wiss.* **1**, 217, de Séguier, *Groupes abstraits*, S. 66.)

Läßt man die Elemente einer Gruppe $\mathfrak{G}$ einfach isomorph sich selbst in anderer Reihenfolge entsprechen, so bezeichnet man eine solche Zuordnung als einen *Isomorphismus der Gruppe in sich selbst.* Man spricht auch von einem *Automorphismus* (Frobenius, *Sitzungsb. d. Berl. Akad.* (1901), 1324), oder von einem *Holomorphismus* (G. A. Miller, *Bull. Am. M. S.* (2) **9**, 112 (1902)). Ein Automorphismus einer Gruppe $\mathfrak{G}$ entsteht beispielsweise, wenn jedem Element G_i aus $\mathfrak{G}$ das Element $B^{-1}G_iB$ zugeordnet wird, wobei B irgendein festes, nicht invariantes Element aus $\mathfrak{G}$ bedeutet. Ein Automorphismus, der sich auf die eben geschilderte Art erzeugen läßt, heißt ein *innerer Automorphismus* oder ein *kogredienter Isomorphismus.* Existieren noch andere Automorphismen, so werden sie als *äußere Automorphismen* oder *kontragrediente Isomorphismen* bezeichnet.

Die Gesamtheit der Automorphismen einer Gruppe hat selbst Gruppencharakter; diese Gruppe heißt die *Isomorphismengruppe* (i-group bei englischen Autoren). *Die inneren Automorphismen bilden eine invariante Untergruppe der Isomorphismengruppe. Ist $\mathfrak{C}$ die Zentrale von $\mathfrak{G}$, so ist die Gruppe der inneren Automorphismen von $\mathfrak{G}$ holoedrisch isomorph mit der Gruppe $\mathfrak{G}/\mathfrak{C}$.*

Sind $G_1, G_2, G_3, \ldots$ die Elemente einer Gruppe $\mathfrak{G}$ und entspricht G_i bei einem Automorphismus der Gruppe $\mathfrak{G}$ das Element G_i' von $\mathfrak{G}$, so kann:

$$\begin{pmatrix} G_1 & G_2 & G_3 \ldots \\ G_1' & G_2' & G_3' \ldots \end{pmatrix}$$

als Permutation aufgefaßt werden; sie bezieht sich auf unendlich viele oder auf eine endliche Anzahl von Symbolen, je nachdem $\mathfrak{G}$ eine unendliche oder eine endliche Anzahl von Elementen besitzt. Die Isomorphismengruppe besteht aus der Gesamtheit aller dieser Permutationen.

Die Hervorhebung der Wichtigkeit der Isomorphismengruppe beginnt mit Hölder, *Math. Ann.* **43**, 313 (1893) und E. H. Moore, *Bull. Am. M. S.* (2) **1**, 61 (1895), die voneinander unabhängig auf die Isomorphismengruppe kamen. (Vgl. Moore, ebenda (2) **2**, 33 (1896).) Die Unterscheidung zwischen kogredienten

und kontragredienten Isomorphismen hatte Klein bereits bei dem speziellen Fall der Ikosaedergruppe in seinen *Vorlesungen über das Ikosaeder*, S. 232, gewonnen. Von Lehrbüchern vgl. über die Isomorphismengruppe: Burnside, *Theory of groups*. S. 221ff.

Alle Elemente einer Gruppe $\mathfrak{G}$ können unter Umständen dadurch entstehen, daß man eine endliche Anzahl von Gruppenelementen $S_1, S_2, \ldots, S_q$ auf jede mögliche Weise untereinander komponiert. In diesem Fall sagt man: *die Gruppe $\mathfrak{G}$ wird von den q Elementen $S_1, S_2, \ldots, S_q$ erzeugt.* Die Elemente $S_1, S_2, \ldots S_q$ heißen ein *System erzeugender Elemente der Gruppe.* Enthält das System $S_1, S_2, \ldots, S_q$ keine überflüssigen Elemente, d. h. ist kein Element S_i $(i = 1, 2, \ldots, q)$ durch Komposition aus den $(q-1)$ übrigen S zu erhalten, so heißen $S_1, S_2, \ldots, S_q$ ein *System unabhängiger erzeugender Elemente der Gruppe* $\mathfrak{G}$. Jede Gleichung, die zwischen ausschließlich unabhängigen erzeugenden Elementen von $\mathfrak{G}$ und der Einheit besteht, heißt eine *fundamentale Definitionsgleichung* der Gruppe; sie läßt sich stets in der Form:

$$f(S) = S_{\alpha_1}^{\lambda_1} S_{\alpha_2}^{\lambda_2} \cdots S_{\alpha_\tau}^{\lambda_\tau} = 1$$

voraussetzen; hierbei sind $\lambda_1, \lambda_2, \ldots \lambda_\tau$ ausschließlich positive Zahlen und $\alpha_1, \alpha_2, \ldots, \alpha_\tau$ Zahlen aus der Reihe $1, 2, \ldots, q$, die beliebig häufig auftreten können $(\tau \gtreqless q)$. Die Gesamtheit fundamentaler Definitionsgleichungen, bei denen diejenigen, die aus den anderen folgen, fortgelassen werden können, heißt *ein System definierender Gleichungen* oder die *Gleichungen der Gruppe.* Durch ein System unabhängiger erzeugender Elemente und ein System definierender Gleichungen ist das Gesetz der Multiplikation irgend zweier Gruppenelemente bekannt und hiermit die Gruppe völlig eindeutig definiert. Da jede Gruppe neben jedem Element das reziproke enthalten muß, findet zwischen den erzeugenden Elementen einer Gruppe stets wenigstens eine Definitionsgleichung statt.

Eine zyklische endliche Gruppe der Ordnung n wird durch ein Element S_1 und die Gleichung $S_1^n = 1$ festgelegt, eine unendliche zyklische Gruppe ist durch zwei unabhängige Elemente S_1, S_2 und die Definitionsgleichung $S_1 S_2 = 1$ bestimmt.

Jede endliche Gruppe läßt sich durch eine endliche Anzahl von Elementen erzeugen, nicht aber jede unendliche Gruppe. Unsere gewöhnlichen rationalen Zahlen bilden bei Ausschluß der

Null bezüglich der Multiplikation eine kommutative Gruppe; wie aus dem Euclidschen Satz von der Existenz unendlichvieler Primzahlen folgt, kann diese Gruppe nicht durch eine endliche Anzahl von Elementen erzeugt werden. In der Theorie der linearen homogenen Differentialgleichungen spielen Gruppen, die durch eine endliche Anzahl von Elementen erzeugt werden, eine fundamentale Rolle. Vgl. L. Schlesinger, *Handbuch der Theorie der Differentialgleichungen*, 2_1, Leipzig 1897, S. 11.

Die denkbar einfachste Gruppe, die von q unabhängigen Elementen $S_1, S_2, \ldots, S_q$ erzeugt wird, ist die unendliche Gruppe $\mathfrak{G}$, bei der zwischen den Elementen die einzige Gleichung $S_1 S_2 \ldots S_q = 1$ besteht. $\overline{\mathfrak{G}}$ sei eine Gruppe, die von den q unabhängigen Elementen $\overline{S}_1, \overline{S}_2, \ldots, \overline{S}_q$ erzeugt wird. Befriedigen diese außer der Relation $\overline{S}_1 \overline{S}_2 \ldots \overline{S}_q = 1$ noch das System der definierenden Gleichungen $f_1(\overline{S}) = 1$, $f_2(\overline{S}) = 1, \ldots, f_n(\overline{S}) = 1$, so ist die Gruppe $\overline{\mathfrak{G}}$ mit der Gruppe $\mathfrak{G}/\mathfrak{J}$ holoedrisch isomorph. Die Gruppe $\mathfrak{J}$ entsteht, wenn man alle Operationen $R^{-1} f_i(S) R$ $(i = 1, 2, \ldots, n)$ der Gruppe $\mathfrak{G}$ auf jede Art komponiert; R bedeutet jedes Element aus $\mathfrak{G}$. Diese Untersuchungen stammen von W. Dyck, *Math. Ann.* **20**, 1 (1882) u. **22**, 70 (1883), ihnen parallel gehen geometrische Darstellungen einer Gruppe durch Einteilung von Bereichen in gleichartige Gebiete. Von Lehrbüchern vgl. W. Burnside, *Theory of groups*, S. 255 bis 310.

§ 3. Abstrakte endliche Gruppen.

Besonders weitreichend werden viele Sätze des vorigen Paragraphen, wenn man sie auf endliche Gruppen anwendet. Um die Ordnung g einer endlichen Gruppe $\mathfrak{G}$ zum Ausdruck zu bringen, schreibt man $\mathfrak{G}_g$.

Fundamentalsatz für endliche Gruppen ist der *Satz von Lagrange* (*Réflexions sur la résolution algébrique des équations* (1770), *Œuvres* **3**): *Jede Untergruppe $\mathfrak{H}$ einer endlichen Gruppe $\mathfrak{G}$ ist selbst eine endliche Gruppe und die Ordnung h von $\mathfrak{H}$ ist ein Divisor der Ordnung g von $\mathfrak{G}$. Der Quotient $j = \frac{g}{h}$*, den man mit $(\mathfrak{G}, \mathfrak{H})$ *bezeichnet, heißt der Index der Untergruppe $\mathfrak{H}$ von $\mathfrak{G}$.*

Der Satz von Lagrange ist eine unmittelbare Folge der Zerlegung der Gruppe $\mathfrak{G}$ in ein System von Nebengruppen. (Vgl. Gleichung (1′) auf S. 183.) Aus dem Satz von Lagrange folgt: *Eine endliche Gruppe enthält nur Elemente von endlicher Ord-*

nung, und die Ordnung jedes Elementes ist ein Divisor der Gruppenordnung.

Bei einer endlichen Gruppe $\mathfrak{G}$ mit den Untergruppen $\mathfrak{H}_1$ und $\mathfrak{H}_2$ erreicht die Zerlegung von $\mathfrak{G}$ nach dem Doppelmodul $(\mathfrak{H}_1, \mathfrak{H}_2)$ stets ihr Ende. Die Gleichung (1) auf S. 182 nimmt die Form an: $\mathfrak{G} = \mathfrak{H}_1 G_1 \mathfrak{H}_2 + \mathfrak{H}_1 G_2 \mathfrak{H}_2 + \cdots + \mathfrak{H}_1 G_t \mathfrak{H}_2$. Es wird: $(\mathfrak{G}, \mathfrak{H}_1) = (\mathfrak{H}_2, \mathfrak{D}_1) + (\mathfrak{H}_2, \mathfrak{D}_2) + \cdots + (\mathfrak{H}_2, \mathfrak{D}_t)$. $(\mathfrak{G}, \mathfrak{H}_1)$ ist der Index $\frac{g}{h_1}$ der Untergruppe $\mathfrak{H}_1$ von $\mathfrak{G}$; $(\mathfrak{H}_2, \mathfrak{D}_i)$ ist der Index $\frac{h_2}{d_i}$ der Untergruppe $\mathfrak{D}_i$ von $\mathfrak{H}_2$; hierbei ist $\mathfrak{D}_i$ $(i = 1, 2, \ldots, t)$ der Durchschnitt der zwei Gruppen $\mathfrak{H}_2$ und $G_i^{-1} \mathfrak{H}_1 G_i$. (Satz von Frobenius, *Journ. f. Math.* **101**, 281 (1887), *Sitzungsb. d. Berl. Akad.* (1895), 167, H. Weber, *Algebra* **2**, 23.)

Auch die Anzahl der Klassen konjugierter Elemente einer endlichen Gruppe $\mathfrak{G}$ der Ordnung g ist endlich. Zerfallen die Elemente von $\mathfrak{G}$ in k Klassen konjugierter Elemente, so mögen die einzelnen Klassen $\nu_1, \nu_2, \ldots, \nu_k$ verschiedene Elemente enthalten; dann ist $g = \nu_1 + \nu_2 + \cdots + \nu_k$. Sind die Zahlen $\nu_1, \nu_2, \ldots, \nu_k$ in der Folge $\nu_1 \leqq \nu_2 \leqq \cdots \leqq \nu_k$ geordnet, so ist $\nu_1 = 1$, da das Einheitselement als invariantes Element für sich allein eine Klasse bildet. Ist A_i ein Element der i^{ten} Klasse, so hat die Gruppe $\mathfrak{B}_i$, die aus den mit A_i vertauschbaren Elementen von $\mathfrak{G}$ besteht, die Ordnung $\frac{g}{\nu_i}$. Die Klasse der mit A_i ähnlichen Elemente besteht aus den ν_i Elementen: $A_i, G_2 A_i G_2^{-1}, G_3 A_i G_3^{-1}, \ldots G_{\nu_i} A_i G_{\nu_i}^{-1}$, wobei die Elemente $G_1 = 1, G_2, G_3, \ldots, G_{\nu_i}$ ein vollständiges Restsystem der Gruppe $\mathfrak{G}$ mod $\mathfrak{B}_i$ vorstellen. Mit Hilfe der angegebenen Resultate beweist man einen der wichtigsten Sätze der Theorie der endlichen Gruppen, nämlich den *Satz von Sylow* (*Math. Ann.* **5**, 586 (1872), eine Ableitung, die von Permutationen keinen Gebrauch macht, hat zuerst Frobenius, *Journ. f. Math.* **100**, 179 (1887) gegeben, ferner ebenda **101**, 282 (1887)):

Ist $\mathfrak{G}$ eine Gruppe der Ordnung g und die Zahl g durch die l^{te} Potenz der Primzahl p teilbar, so besitzt $\mathfrak{G}$ stets eine Untergruppe der Ordnung p^l.

In dem speziellen Fall $l = 1$ findet sich dieses Theorem bereits bei Cauchy in den *Exercices d'analyse* **3**, 250 (1844).

Ist p^m die *höchste* Potenz einer Primzahl p, die in g enthalten ist, so bezeichnet man die Untergruppen der Ordnung p^m einer Gruppe $\mathfrak{G}$ der Ordnung g als *Sylowsche Untergruppen* von $\mathfrak{G}$ (G. A. Miller, *Bull. Am. M. S.* (2) **9**, 543 (1903)).

Von seinen Untergruppen hat Sylow bewiesen: *Alle Sylowschen Untergruppen der Ordnung p^m einer Gruppe $\mathfrak{G}$ sind miteinander innerhalb $\mathfrak{G}$ ähnlich; ihre Anzahl ϱ ist eine Zahl der Form $p\varrho' + 1$, also kongruent 1 (mod p). Ist $\mathfrak{S}$ irgendeine Sylowsche Untergruppe von $\mathfrak{G}$ der Ordnung p^m, so hat $\mathfrak{G}$ die Ordnung $p^m \cdot \varrho \cdot r$; hierbei ist r nicht durch p teilbar und $p^m \cdot r$ die Ordnung der größten Untergruppe von $\mathfrak{G}$, deren Elemente mit der Gruppe $\mathfrak{S}$ vertauschbar sind.*

Jede Untergruppe von $\mathfrak{G}$ der Ordnung p^l ist in einer der ϱ ähnlichen Sylowschen Untergruppen von $\mathfrak{G}$ der Ordnung p^m enthalten.

Den von Sylow nur für die Anzahl ϱ der Sylowschen Untergruppen einer Gruppe $\mathfrak{G}$ bewiesenen Satz hat Frobenius (*Sitzungsb. d. Berl. Akad.* (1895), 981) dahin erweitert: *Die Anzahl der Untergruppen von $\mathfrak{G}$ der Ordnung p^l, wobei g auch durch eine höhere Potenz von p als die l^{te} teilbar sein kann, ist stets $\equiv 1$ (mod. p). Die Gruppen der Ordnung p^l sind, wenn $l < m$ ist, nicht notwendig miteinander ähnlich.*

Über die Sylowschen Sätze vgl. man noch Frobenius, *Sitzungsb. d. Berl. Akad.* (1895), S. 170, sowie von Lehrbüchern: H. Weber, *Algebra* **2**, 135, W. Burnside, *Theory of groups*, S. 90, de Séguier, *Groupes abstraits*, S. 79, ferner G. A. Miller, *Bull. Am. M. S.* (2) **14**, 91 (1907), *Proc. Lond. M. S.* (2) **2**, 142 (1905).

Im Zusammenhang mit seiner Erweiterung des Sylowschen Satzes gelangte Frobenius (*Sitzungsb. d. Berl. Akad.* (1895), 981, ebenda (1903), 987) zu folgendem Fundamentalsatz über *die Potenzen eines Elementes einer endlichen Gruppe:*

Die Anzahl t der Elemente X einer Gruppe $\mathfrak{G}_g$, deren n^{te} Potenz, also X^n, gleich einem beliebig vorgegebenen Element A von $\mathfrak{G}$ ist, ist durch den größten gemeinsamen Divisor d von n und v teilbar, wenn v die Anzahl der mit A vertauschbaren Gruppenelemente, also $\frac{g}{v}$ die Zahl der mit A ähnlichen Gruppenelemente von $\mathfrak{G}$ ist. Die Zahl t kann auch den Wert Null haben.

Ist A ein *invariantes Element*, also im besonderen das Einheitselement E, so ist v gleich der Gruppenordnung g. Ist g durch n teilbar, so wird für ein invariantes Element die Zahl $d = n$. Die Gleichung $X^n = E$ wird stets durch $X = E$ befriedigt; daher ist für diesen Fall $t \neq 0$. Folglich ergibt sich:

Ist die Ordnung g einer Gruppe $\mathfrak{G}$ durch die Zahl n teilbar, so ist die Anzahl t derjenigen Elemente der Gruppe $\mathfrak{G}$, deren

Ordnung in n aufgeht, ein Vielfaches von n. Ist die Ordnung einer Gruppe $\mathfrak{G}$ durch n teilbar, so ist die von den t Elementen, welche der Gleichung $X^n = E$ genügen, erzeugte Gruppe entweder die Gruppe $\mathfrak{G}$ oder eine charakteristische Untergruppe von $\mathfrak{G}$, deren Ordnung ebenso wie die Zahl t durch n teilbar ist.

Von Lehrbüchern vgl. Burnside, *Theory of groups*, S. 110, de Séguier, *Groupes abstraits*, S. 74.

Anknüpfend an den Begriff der größten invarianten Untergruppe und die Sätze auf S. 188 erhält man für endliche Gruppen besonders bemerkenswerte Sätze: Ist $\mathfrak{G}$ eine endliche Gruppe, so heißt eine mit der Einheit endende Reihe von Gruppen: $\mathfrak{G}$, $\mathfrak{G}_1$, $\mathfrak{G}_2$, $\mathfrak{G}_3$, ..., 1, von denen jede Gruppe $\mathfrak{G}_i$ $(i = 1, 2, 3, \ldots)$ eine größte invariante Untergruppe der voraufgehenden $\mathfrak{G}_{i-1}$ ist, eine *Kompositionsreihe* oder eine *Reihe der Zusammensetzung der Gruppe* $\mathfrak{G}$. Für die Faktorgruppen $\mathfrak{G}/\mathfrak{G}_1$, $\mathfrak{G}_1/\mathfrak{G}_2$, $\mathfrak{G}_2/\mathfrak{G}_3$, ... beweist man: *Besitzt eine endliche Gruppe zwei verschiedene Kompositionsreihen, so sind, von der Reihenfolge abgesehen, die sich bei der einen Reihe ergebenden Faktorgruppen den aus der anderen Reihe entspringenden holoedrisch isomorph* (Hölder, *Math. Ann.* **34**, 37 (1889)). Sind $g, g_1, g_2, \ldots, 1$ die Ordnungen der Gruppen $\mathfrak{G}$, $\mathfrak{G}_1$, $\mathfrak{G}_2$, ..., 1, so heißen die Zahlen $\frac{g}{g_1} = e_1$, $\frac{g_1}{g_2} = e_2$, ... die *numerischen Faktoren der Zusammensetzung* oder die *Indexreihe der Gruppe* $\mathfrak{G}$. Als Ordnungen der Faktorgruppen $\mathfrak{G}/\mathfrak{G}_1$, $\mathfrak{G}_1/\mathfrak{G}_2$, ... sind, von der Reihenfolge abgesehen, die Zahlen $e_1, e_2, \ldots$ *eindeutig* bestimmt (C. Jordan, *Traité*, p. 41). Neben der Reihe der Zusammensetzung kann man für eine endliche Gruppe $\mathfrak{G}$ auch eine *Hauptreihe* definieren (C. Jordan, *Traité*, p. 663). Die Gruppen: $\mathfrak{G}$, $\mathfrak{H}_1$, $\mathfrak{H}_2$, ..., 1 bilden eine Hauptreihe, wenn jede Gruppe $\mathfrak{H}_i$ $(i = 1, 2, 3, \ldots)$ nicht nur eine invariante Untergruppe der vorangehenden $\mathfrak{H}_{i-1}$ $(\mathfrak{H}_0 = \mathfrak{G})$, sondern sogar in der Gesamtgruppe $\mathfrak{G}$ invariant ist und zwischen zwei aufeinanderfolgenden Gruppen $\mathfrak{H}_{i-1}$ und $\mathfrak{H}_i$ es nie möglich ist, eine neue Gruppe einzuschieben, ohne daß die Reihe ihren Charakter verliert.

Wie auch immer die Hauptreihe $\mathfrak{G}$, $\mathfrak{H}_1$, $\mathfrak{H}_2$, ..., 1 einer endlichen abstrakten Gruppe gewählt sein mag, stets sind die Faktorgruppen $\mathfrak{G}/\mathfrak{H}_1$, $\mathfrak{H}_1/\mathfrak{H}_2$, $\mathfrak{H}_2/\mathfrak{H}_3$, ... bis auf die Reihenfolge dieselben, wenn man holoedrisch isomorphe Gruppen als nicht verschieden ansieht (Hölder, *Math. Ann.* **34**, 38 (1889)); daß die Zahlenfaktoren $\frac{g}{h_1}, \frac{h_1}{h_2}, \frac{h_2}{h_3}, \ldots$, abgesehen von der Reihenfolge, eindeutig bestimmt sind, bereits bei C. Jordan, a. a. O.). Jede

der aus einer Hauptreihe hervorgehenden Faktorgruppen $\mathfrak{G}/\mathfrak{H}_1$, $\mathfrak{H}_1/\mathfrak{H}_2$, $\mathfrak{H}_2/\mathfrak{H}_3$, ... ist eine elementare Gruppe (Hölder, a. a. O.). Eine Gruppe heißt nach Frobenius (*Sitzungsb. d. Berl. Akad.* (1902), 358) *elementar*, wenn sie einfach oder das direkte Produkt mehrerer holoedrisch isomorpher einfacher Gruppen ist. Das charakteristische Kennzeichen einer elementaren Gruppe ist, daß sie keine charakteristische Untergruppe besitzt.

Durch Einschieben von Gruppen kann man aus einer Hauptreihe eine Kompositionsreihe konstruieren, hingegen geht nicht stets durch Unterdrücken von Gruppen aus letzterer eine Hauptreihe hervor. Zu einer endlichen Gruppe $\mathfrak{G}$ kann man stets eine derartige Kompositionsreihe $\mathfrak{G}$, $\mathfrak{G}_1$, $\mathfrak{G}_2$, ..., $\mathfrak{G}_i$, ..., 1 finden, daß, so oft sich die numerischen Faktoren der Zusammensetzung ändern, also $\frac{g_{i-1}}{g_i}$ von $\frac{g_i}{g_{i+1}}$ verschieden ist, $\mathfrak{G}_i$ eine invariante Untergruppe von $\mathfrak{G}$ ist (vgl. H. Weber, *Algebra* **2**, 33). Zu einer endlichen Gruppe läßt sich auch eine *lückenlose Reihe charakteristischer Untergruppen* konstruieren (Frobenius, *Sitzungsb. d. Berl. Akad.* (1895), 1027).

Ist g eine gegebene Zahl, so existiert nur eine endliche Anzahl nicht isomorpher abstrakter endlicher Gruppen, welche die Ordnung g besitzen. Man kann sich daher die Aufgabe stellen, alle nicht isomorphen abstrakten endlichen Gruppen der vorgegebenen Ordnung g mit Angabe ihrer definierenden Gleichungen aufzuzählen. Ist p eine Primzahl, so gibt es nur eine Gruppe p^{ter} Ordnung, nämlich die durch $A^p = 1$ definierte zyklische Gruppe. Nicht isomorphe Gruppen der Ordnung p^2 existieren nur 2, die beide Abelsche Gruppen sind; solche der Ordnung p^3 gibt es 3 Abelsche und 2 nicht Abelsche (Hölder, *Math. Ann.* **43**, 301 (1893)). Sind p und q voneinander verschiedene Primzahlen ($p > q$), so gibt es, wenn $p - 1$ nicht durch q teilbar ist, nur eine Gruppe der Ordnung pq, diese ist zyklisch. Im anderen Fall gibt es noch eine zweite Gruppe der Ordnung pq; diese wird durch $A_1^p = 1$, $A_2^q = 1$, $A_2^{q-1} A_1 A_2 A_1^{p-\alpha} = 1$ definiert, wobei α eine beliebige Zahl bedeutet, die mod p zum Exponenten q gehört (Netto, *Substitutionentheorie*, S. 134, Hölder, a. a. O., S. 312). Zusammenfassende Angaben über die verschiedenen Typen nicht isomorpher abstrakter endlicher Gruppen bei Easton, *The constructive development of group-theory*, S. 77, vgl. zur Ergänzung: G. A. Miller, *Bull. Am. M. S.* (2) **14**, 89 (1907).

Ist die Zahl k der Klassen konjugierter Elemente einer endlichen abstrakten Gruppe gegeben, so gibt es nur eine endliche Anzahl nicht isomorpher endlicher abstrakter Gruppen, welche die vorgegebene Klassenzahl k besitzen (Landau, *Math. Ann.* **56**, 674 (1903)).

Die Isomorphismengruppe einer endlichen Gruppe der Ordnung g ist höchstens von der Ordnung $(g-1)!$ und erreicht diesen Maximalwert nur für $g = 1, 2, 3$ und die nicht zyklische Gruppe der Ordnung 4. Eine Besprechung der Untersuchungen über die Isomorphismengruppe bei G. A. Miller, *Bull. Am. M.* S. (2) **14**, 124 (1907).

§ 4. Auflösbare, einfache und zusammengesetzte Gruppen.

Eine endliche Gruppe heißt auflösbar (C. Jordan), wenn ihre Indexreihe aus lauter Primzahlen besteht. H. Weber (*Algebra* **1**, 647) bezeichnet die auflösbaren Gruppen als *metazyklisch.*

Eine Gruppe ist dann und nur dann auflösbar, wenn die Ordnungen der Faktorgruppen, die irgendeine Hauptreihe der Gruppe liefert, Primzahlen oder ihre Potenzen sind.

Eine Gruppe ist dann und nur dann auflösbar, wenn für ihre sukzessiven Kommutatorgruppen folgendes eintritt: Bildet man die Kommutatorgruppe der Gruppe, von dieser wiederum die Kommutatorgruppe usw., so gelangt man schließlich zur Einheit (C. Jordan, *Traité,* S. 395; G. A. Miller, *Quart. J.* **28**, 268 (1896)).

Alle Gruppen der Ordnung p^m (p Primzahl) sind stets auflösbar (Sylow, *Math. Ann.* **5**, 588 (1872), Frobenius, *Sitzungsb. d. Berl. Akad.* (1895), 173 u. 982, W. Burnside, *Theory of groups,* Chap. V, de Séguier, *Groupes abstraits,* Chap. IV.)

Die Gruppen der Ordnung p^m und diejenigen, die das direkte Produkt von ihnen sind, erschöpfen sämtliche Gruppen, bei denen sich *jede* Untergruppe in eine der verschiedenen Kompositionsreihen, die man zu der gegebenen Gruppe konstruieren kann, einordnen läßt. Solche Gruppen heißen *spezielle Gruppen.* (W. Burnside, *Theory of groups,* S. 115, A. Loewy, *Math. Ann.* **55**, 67 (1902), de Séguier, *Groupes abstraits,* S. 87).

Satz von Frobenius (*Sitzungsb. d. Berl. Akad.* (1893), 342, (1895), 1043, Hölder, *Gött. Nachr.* (1895), 211, H. Weber, *Algebra* **2**, 140): *Jede Gruppe, deren Ordnung ein Produkt lauter verschiedener Primzahlen ist, ist auflösbar.*

Es gilt folgender allgemeiner Satz (Frobenius, *Sitzungsb. d. Berl. Akad.* (1895), 1041): *Ist* $\mathfrak{G}$ *eine Gruppe der Ordnung* $g = p_1^{\lambda_1} p_2^{\lambda_2} \dots p_n^{\lambda_n}$, *wobei* $p_1 < p_2 < \dots < p_n$ *ihrer Größe nach geordnete Primzahlen bedeuten, und sind sämtliche in* $\mathfrak{G}$ *enthaltene Sylowsche Gruppen* $\mathfrak{P}_i$ *der Ordnungen* $p_i^{\lambda_i}$ $(i = 1, 2, \dots, n-1)$ *ausnahmslos kommutative Gruppen vom Range* 1 *oder* 2 (vgl. S. 203), *so ist* $\mathfrak{G}$ *eine auflösbare Gruppe. Eine Ausnahme kann nur eintreten, wenn* $p_1 = 2$, $p_2 = 3$, $\mathfrak{P}_1$ *den Rang* 2 *hat und* $\mathfrak{G}$ *eine Untergruppe besitzt, deren Ordnung in* $2^{\lambda_1} 3^{\lambda_2}$ *aufgeht und mit der Tetraëdergruppe der Ordnung* 12 *isomorph ist.*

Haben die Exponenten $\lambda_1, \lambda_2, \dots, \lambda_{n-1}$ die Werte 1 oder 2, so sind die Gruppen $\mathfrak{P}_i$ $(i = 1, 2, \dots, n-1)$ sicher kommutative Gruppen vom Range 1 oder 2. Mithin folgt: *Jede Gruppe der Ordnung* $p_1^{\lambda_1} p_2^{\lambda_2} \dots p_n^{\lambda_n}$, *bei der* $p_1 < p_2 < \dots < p_n$ *ihrer Größe nach geordnete Primzahlen sind, jeder der Exponenten* $\lambda_1, \lambda_2, \dots, \lambda_{n-1}$ *den Wert* 1 *oder* 2 *besitzt,* λ_n *eine beliebige ganze positive Zahl bedeutet, ist auflösbar; eine Ausnahme kann nur* $p_1 = 2$, $p_2 = 3$, $\lambda_1 = 2$ *bilden.*

Satz von W. Burnside (*Proc. Lond. M. S.* (2) **1**, 388 (1904) u. (2) **2**, 432 (1905)): *Die Gruppen der Ordnung* $p^\alpha q^\beta$ (*p und q verschiedene Primzahlen*) *sind auflösbar.* (Spezialfälle sind früher von Frobenius, Burnside, Jordan behandelt worden, vgl. H. Weber, *Algebra* **2**, 145 und Frobenius, *Acta math.* **26**, 189 (1902), siehe auch Cole, *Trans. Am. M. S.* **5**, 214 (1904)). Der Beweis des Burnsideschen Satzes ergibt sich aus seinem (*Proc. Lond. M. S.* (2) **1**, 392) Theorem: *Jede endliche Gruppe, bei der die Anzahl von Elementen in einer Klasse konjugierter Elemente eine Primzahlpotenz ist, kann nicht einfach sein.*

Jede Gruppe, deren Ordnung das Produkt von drei Primzahlen ist, ist auflösbar. Über Gruppen, deren Ordnungen das Produkt von 4 und 5 Primzahlen sind, vgl. unten.

Jede Gruppe der Ordnung $p^4 qr$, *wobei* p, q, r *beliebige verschiedene ungerade Primzahlen bedeuten, ist auflösbar* (Burnside, *Proc. Lond. M. S.* **33**, 266 (1901), Frobenius, *Sitzungsb. d. Berl. Akad.* (1901), 1329).

Folgende Sätze von Frobenius erweisen eine Gruppe als zusammengesetzt:

Ist die Gruppe $\mathfrak{H}_h$ *in einer Gruppe* $\mathfrak{G}$ *der Ordnung* hn *enthalten und ist sie darin mit* n *verschiedenen Gruppen konjugiert, von denen je zwei teilerfremd sind, so enthält* $\mathfrak{G}$ *eine und nur eine, demnach charakteristische Untergruppe der Ordnung* n (*Sitzungsb. d. Berl. Akad.* (1901), 1226).

Enthält die Ordnung g einer Gruppe $\mathfrak{G}$ die Primzahl p in der ersten Potenz und ist $p-1$ und g teilerfremd, so enthält $\mathfrak{G}$ eine und auch nur eine, demnach charakteristische Untergruppe der Ordnung $\frac{g}{p}$ (*Sitzungsb. d. Berl. Akad.* (1901), 849).

Verallgemeinerungen ebenda, vgl. ferner J. Schur, *Sitzungsb. d. Berl. Akad.* (1902), 1013.

Einfache, nicht isomorphe Gruppen, deren Ordnungen zusammengesetzt und kleiner als 2000 sind, gibt es nur sechs, nämlich je eine der Ordnungen: 60, 168, 360, 504, 660 und 1092 (Galois (1—60), *Œuvres*, p. 26, Hölder, *die einfachen Gruppen im ersten und zweiten Hundert der Ordnungszahlen*, *Math. Ann.* **40**, 55 (1892), Cole (200—600), *Am. J. math.* **14**, 378 (1892) und **15**, 303 (1893), W. Burnside (660—1092), *Proc. Lond. M.* S. **26**, 325 (1895), Ling and G. A. Miller (1092—2001), *Am. J. math.* **22**, 13 (1900)). Diese sechs Gruppen sind mit Gruppen des unendlichen Systems einfacher Gruppen holoedrisch isomorph, das die *verallgemeinerte endliche Modulargruppe* der Ordnung $\frac{p^m(p^{2m}-1)}{d}$ ($d = 1$ oder 2, je nachdem die Primzahl $p = 2$ oder größer als 2 ist) für die Werte $p^m = 5$ (oder 2^2), 7, 3^2, 2^3, 11, 13 (vgl. § 7) liefert. Die einfache Gruppe der Ordnung 60 wird durch zwei unabhängige Elemente S_1 und S_2 erzeugt; $S_1^5 = 1$, $S_2^2 = 1$, $(S_1 S_2)^3 = 1$ sind die Definitionsgleichungen der Gruppe.

Unter allen Gruppen, deren Ordnungen das Produkt von 4 oder 5 Primzahlen sind, sind nur die vier Gruppen der Ordnungen 60, 168, 660 *und* 1092 einfach (Frobenius, *Sitzungsb. d. Berl. Akad.* (1895), 1041, über Gruppen, deren Ordnungen das Produkt von 4 Primzahlen ist, vgl. auch Glenn, *Trans. Am. M. S.* **7**, 137 (1906)).

Ein Verzeichnis der bekannten einfachen Gruppen gibt Dickson, *Linear groups*, S. 309. In dieser Liste nicht aufgeführt ist das später bekannt gewordene unendliche System einfacher Gruppen der Ordnung $p^{6q}(p^{6q}-1)(p^{2q}-1)$, das für jede Primzahl $p > 2$ und jede ganze Zahl $q \geqq 1$ und für $p = 2$ und jedes ganzzahlige $q > 1$ existiert (Dickson, *Trans. Am. M. S.* **2**, 389 (1901), *Math. Ann.* **60**, 137 (1905)). Für jede Zahl der Form $\frac{1}{2}(p^{m(2r)}-1) \cdot p^{m(2r-1)} \cdot (p^{m(2r-2)}-1)$ $p^{m(2r-3)} \cdots (p^{2m}-1)p^m$, wobei p eine beliebige, von 2 verschiedene Primzahl, r und m beliebige ganze positive Zahlen,

$r > 2$ bedeuten, existieren sogar zwei einfache, nicht isomorphe Gruppen der angegebenen Ordnungszahl (Dickson, *Linear groups*, S. 309). Abgesehen von der abstrakten Gruppe der Ordnung $\frac{n!}{2}$, die mit der alternierenden Permutationsgruppe (vgl. § 6) isomorph ist, entstammen die bekannten Systeme einfacher Gruppen der Theorie der Kongruenzgruppen (vgl. § 12); die einfache Gruppe niedrigster Ordnung, die sich keinem System einordnen läßt, ist von der Ordnung 7920; sie läßt sich als vierfach transitive Permutationsgruppe des Grades 11 darstellen (Cole, *Quart. Journ.* **27**, 48 (1895), de Séguier, *Journ. de math.* (5) **8**, 291 (1902)). Die weiteren bekannten einfachen Gruppen, die sich bisher keinem System einordnen lassen, sind zwei Gruppen der Ordnungen 95040 und 244823040, die mit den zwei von Mathieu (*Journ. de math.* (2) **6**, 270 (1861), (2) **18**, 25 (1873)) entdeckten fünffach transitiven Permutationsgruppen der Grade 12 und 24 (W. Burnside, *Theory of groups*, 220, Frobenius, *Sitzungsb. d. Berl. Akad.* (1904), 567) holoedrisch isomorph sind, sowie zwei größte Untergruppen der letzten Gruppe, die als Permutationsgruppen der Grade 22 und 23 darstellbar sind (G. A. Miller, *Bull. de la soc. math.* **28**, 266 (1900)). Eine einfache Gruppe ungerader, zusammengesetzter Ordnung ist nicht bekannt. (Über den Stand dieser Frage vgl. Rietz, *Am. J. math.* **26**, 1 (1904).)

Nicht auflösbare abstrakte Gruppen, deren Ordnungen < 480 ist, gibt es nur die folgenden 25 nicht isomorphen:

Ordnung:	60	120	168	180	240	300	336	360	420
Anzahl:	1	3	1	1	8	1	3	6	1

(Hölder, *Math. Ann.* **46**, 420 (1895)).

§ 5. Abelsche Gruppen. Hamiltonsche Gruppen. Die Quaternionengruppe.

Jede Gruppe, die nur aus vertauschbaren Operationen besteht, heißt eine *Abelsche* oder *kommutative* Gruppe (vgl. S. 174). In jeder endlichen abstrakten Abelschen Gruppe $\mathfrak{G}$ kann man stets ein System erzeugender Elemente $A_1, A_2, \ldots, A_\sigma$ der Ordnungen $g_1, g_2, \ldots, g_\sigma$ derartig auswählen, daß das Produkt $g_1 g_2 \ldots g_\sigma$ gleich der Ordnung von $\mathfrak{G}$ ist und jedes Element von $\mathfrak{G}$ ein und auch nur einmal in der Form $A_1^{h_1} A_2^{h_2} \ldots A_\sigma^{h_\sigma}$ erscheint, wenn $h_1, h_2, \ldots, h_\sigma$ alle ganzzahligen Werte von 0 bis

$g_1 - 1$, bzw. 0 bis $g_2 - 1$ usw., 0 bis $g_\sigma - 1$ annehmen. Ein System von Elementen, das sich zu einer solchen Darstellung eignet, heißt eine *Basis der Abelschen Gruppe;* die erzeugenden Elemente der Basis heißen *Basiselemente.*

Eine Basis einer Abelschen Gruppe kann verschiedenartig gewählt werden; auch die Anzahl der Basiselemente ist nicht für jede Gruppe die gleiche. Bei jeder endlichen Abelschen lassen sich die Basiselemente derartig wählen, daß ihre Ordnungen Primzahlen oder ihre Potenzen sind. Die auf diese Weise gewonnenen Primzahlpotenzen heißen die *Invarianten der Gruppe.* Zwei Abelsche Gruppen sind dann und nur dann holoedrisch isomorph, wenn sie die gleichen Invarianten haben.

Eine Abelsche Gruppe von Primzahlpotenzordnung $p^m (m \geqq 1)$ mit den Invarianten $p^{\alpha_1}, p^{\alpha_2}, \ldots p^{\alpha_l}$ $(\alpha_1 + \alpha_2 + \cdots + \alpha_l = m)$ heißt vom *Typus* $(\alpha_1, \alpha_2, \ldots, \alpha_l)$. Da alle Abelschen Gruppen mit denselben Invarianten isomorph sind, existieren bei gegebenem p und m nur soviel verschiedene Abelsche Gruppen der Ordnung p^m, als es verschiedene additive Zerlegungen der Zahl m gibt.

Bei jeder endlichen Abelschen Gruppe kann man eine Basis auch derartig wählen, daß die Ordnungszahl jedes voraufgehenden Basiselementes A_i entweder durch die Ordnungszahl des folgenden A_{i+1} teilbar oder ihr gleich ist. Auf diese Weise erhält man die kleinste Anzahl von Basiselementen; diese kleinste Anzahl erzeugender Elemente der Abelschen Gruppe heißt der *Rang* der Abelschen Gruppe. Der Begriff der Invarianten wurde allerdings in anderer Fassung von Frobenius u. Stickelberger (*Journ. f. Math.* **86**, 236 (1879)) eingeführt; die obige Definition geht auf H. Weber zurück. Vgl. H. Weber, *Algebra* **2**, 38ff. Ist m eine beliebige ganze Zahl, so bilden die $\varphi(m)$ ganzen positiven Zahlen, die kleiner als m und relativ prim zu m sind, wenn man sie multiplikativ verknüpft und die sich bei der Multiplikation ergebenden Zahlen stets mod m nimmt, eine Abelsche Gruppe der Ordnung $\varphi(m)$.

Wegen unendlicher Abelscher Gruppen vgl. H. Weber, *Math. Ann.* **48**, 435 (1897) und de Séguier, *Groupes abstraits,* S. 97.

Jede endliche Gruppe, deren sämtliche Untergruppen invariant sind, heißt eine *Hamiltonsche Gruppe.* (Dedekind, *Math. Ann.* **48**, 548 (1897), G. A. Miller, *C. R.* **126**, 1406 (1898),

Bull. Am. M. S. (2) **4**, 510, (1898), E. Wendt, *Math. Ann.* **59**, 187 (1904), **60**, 319 (1905)). Eine besonders wichtige Hamiltonsche Gruppe ist die *Quaternionengruppe;* hierunter versteht man die Gruppe 8ter Ordnung, die durch die zwei unabhängigen Elemente i_1 und i_2 erzeugt wird und deren definierende Gleichungen: $i_1^4 = 1$, $i_1^2 i_2^2 = 1$, $i_1 i_2 i_1 i_2^3 = 1$ lauten. Für die Zusammensetzung der Quaternioneneinheiten (vgl. S. 16 und setze $i_1 i_2 = i_3$ und $i_1^2 = -1$) gelten genau die gleichen Kompositionsregeln wie für die Quaternionengruppe. Die Quaternionengruppe besitzt drei Untergruppen vierter und eine zweiter Ordnung.

Jede Hamiltonsche Gruppe, die keine Abelsche Gruppe ist, ist das direkte Produkt einer Quaternionengruppe, einer Abelschen Gruppe, deren sämtliche Elemente die Ordnung 2 haben, und einer Abelschen Gruppe ungerader Ordnungszahl. Umgekehrt ist jede derartige Gruppe eine Hamiltonsche Gruppe. Verallgemeinerungen der Hamiltonschen Gruppen: G. A. Miller, *Math. Ann.* **60**, 597 (1905), *Arch. f. Math.* (3) **11**, 76 (1906), Wendt, *Math. Ann.* **62**, 381 (1906).

Jede Untergruppe einer Abelschen Gruppe ist wiederum Abelsch. Es gibt auch nicht-Abelsche Gruppen, deren sämtliche Untergrupppen Abelsche Gruppen sind; ihre Ordnung ist nie durch mehr als zwei verschiedene Primzahlen teilbar (Miller and Moreno, *Trans. Am. M. S.* **4**, 398 (1903)). Eine Gruppe, deren Kommutatorgruppe nur aus invarianten Elementen besteht, heißt *metabelsch* (W. B. Fite, *Trans. Am. M. S.* **3**, 331 (1902)).

§ 6. Permutationen. Symmetrische und alternierende Gruppe.

Diejenige Operation, die n Symbole (man sagt auch Ziffern, Zahlen, Buchstaben, Marken) durch die gleichen Symbole in der nämlichen oder einer anderen Anordnung ersetzt, heißt eine *Permutation* oder *Substitution* (in engerem Sinn). Bezeichnet man die zu vertauschenden Symbole mit $1, 2, 3, \ldots, n$ und irgendeine Anordnung von ihnen mit $\alpha_1, \alpha_2, \ldots, \alpha_n$, so schreibt man die Permutation, die 1 durch α_1, 2 durch α_2, $\ldots$, n durch α_n ersetzt, $\begin{pmatrix} 1 & 2 & 3 & \ldots & n \\ \alpha_1 & \alpha_2 & \alpha_3 & \ldots & \alpha_n \end{pmatrix}$. (Cauchy, *Œuvres* (2) **1**, 67, manche Autoren schreiben umgekehrt nach Cauchys späteren Arbeiten, *Œuvres* (1) **9**, 281 $\begin{pmatrix} \alpha_1 & \alpha_2 & \alpha_3 & \ldots & \alpha_n \\ 1 & 2 & 3 & \ldots & n \end{pmatrix}$.) Häufig

bezeichnet man eine Permutation abgekürzt mit einem einzigen Buchstaben A.

Jede Permutation kann auf $n!$ Arten geschrieben werden, indem man entweder den Zähler oder den Nenner beliebig anordnet und nur den Zusammenhang zwischen den übereinanderstehenden Symbolen nicht stört. Ist $a_1, a_2, \ldots, a_n$ irgendeine Anordnung der Zahlen $1, 2, \ldots, n$, so kann man die Permutation $A = \begin{pmatrix} 1 & 2 & 3 & \ldots & n \\ \alpha_1 & \alpha_2 & \alpha_3 & \ldots & \alpha_n \end{pmatrix}$ auch gleichwertig $\begin{pmatrix} a_1 & a_2 & \ldots & a_n \\ \alpha_{a_1} & \alpha_{a_2} & \ldots & \alpha_{a_n} \end{pmatrix}$ schreiben. Bei einer Permutation läßt man häufig die sich nicht ändernden Symbole fort.

Die Anzahl n der Symbole, auf die sich die Permutation bezieht, heißt der *Grad der Permutation*; die Anzahl der Symbole, die durch von ihnen verschiedene ersetzt werden, heißt die *Klasse der Permutation.*

Aus zwei Permutationen

$$A = \begin{pmatrix} 1 & 2 & \ldots & n \\ \alpha_1 & \alpha_2 & \ldots & \alpha_n \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} 1 & 2 & \ldots & n \\ \beta_1 & \beta_2 & \ldots & \beta_n \end{pmatrix} = \begin{pmatrix} \alpha_1 & \alpha_2 & \ldots & \alpha_n \\ \beta_{\alpha_1} & \beta_{\alpha_2} & \ldots & \beta_{\alpha_n} \end{pmatrix}$$

entspringt eine dritte

$$C = \begin{pmatrix} 1 & 2 & \ldots & n \\ \beta_{\alpha_1} & \beta_{\alpha_2} & \ldots & \beta_{\alpha_n} \end{pmatrix},$$

das *Produkt* von A und B. Man schreibt $C = AB$. (Manche Autoren bezeichnen umgekehrt das so gebildete Produkt C mit BA.)

Eine Permutation heißt *zyklisch* oder *zirkular*, wenn sich alle ihre Symbole oder bei Fortlassung der sich nicht ändernden die übrigen so anordnen lassen, daß das erste Symbol durch das zweite, das zweite durch das dritte, usw., das letzte durch das erste ersetzt wird. Eine zyklische Permutation hat bei Fortlassung der unverändert bleibenden Symbole die Form:

$$\begin{pmatrix} \alpha_1 & \alpha_2 & \ldots & \alpha_{m-1} & \alpha_m \\ \alpha_2 & \alpha_3 & \ldots & \alpha_m & \alpha_1 \end{pmatrix} \qquad (m \leqq n).$$

Man bezeichnet sie mit $(\alpha_1, \alpha_2, \ldots, \alpha_m)$, indem man die Symbole in der Reihenfolge, in der das eine für das andere tritt, in eine Klammer hintereinander setzt. Die zyklische Permutation läßt sich daher auf m Arten schreiben:

$$(\alpha_1, \alpha_2, \ldots, \alpha_m) = (\alpha_2, \alpha_3, \ldots, \alpha_m, \alpha_1) = \cdots = (\alpha_m, \alpha_1, \alpha_2, \ldots, \alpha_{m-1}).$$

Eine Permutation, bei der nur zwei Symbole vertauscht

werden, also eine zyklische Permutation (a, b), heißt eine *Transposition.*

Jede Permutation kann, abgesehen von der Reihenfolge, auf eine und auch nur auf eine Weise als ein Produkt zyklischer Permutationen, die kein Symbol gemeinsam haben, dargestellt werden. Diese zyklischen Faktoren heißen die *Zykeln der betr. Permutation.*

Eine zyklische Permutation oder eine solche, bei der jeder Zyklus, abgesehen von den eingliedrigen, die gleiche Anzahl von Symbolen enthält, heißt eine *reguläre Permutation.*

Jede Permutation n^{ten} Grades mit r Zykeln, wobei die eingliedrigen mitzurechnen sind, kann durch $n - r$ Transpositionen ersetzt werden. (Vgl. S. 44.) *Jede Permutation kann auf unendlich viele Weisen als Produkt von Transpositionen dargestellt werden; dabei ist sie entweder das Produkt einer stets geraden oder einer stets ungeraden Anzahl von Transpositionen.*

Eine Permutation heißt *gerade, von der ersten Klasse* oder *eigentlich*, wenn sie aus einer geraden Anzahl von Transpositionen gebildet werden kann; im anderen Fall heißt sie *ungerade*, von der *zweiten Klasse* oder *uneigentlich.*

Ist A eine gegebene Permutation n^{ten} Grades, so ist die Anzahl der Arten, auf die sich A als Produkt von w Transpositionen darstellen läßt, gleich

$$c_1 f_1^{\,w} + c_2 f_2^{\,w} + \cdots + c_k f_k^{\,w};$$

hierbei hängen $f_1, f_2, \ldots, f_k$ nur von n allein ab und sind ganze Zahlen, $c_1, c_2, \ldots, c_k$ sind rationale, von n und der gegebenen Permutation, aber nicht von w abhängige Zahlen, die mit den Gruppencharakteren der symmetrischen Gruppe (vgl. § 10) in Zusammenhang stehen. (A. Hurwitz, *Math. Ann.* **39**, 7 (1891), **55**, 53 (1902), Netto, ebenda **56**, 482 (1903), Frobenius, *Sitzungsb. d. Berl. Akad.* (1903), 357.)

Die $n!$ verschiedenen Permutationen von n Symbolen bilden bei ihrer Komposition eine endliche Gruppe, die symmetrische Gruppe der Ordnung $n!$ (Der Name von C. Jordan, *Journ. de math.* (2) **16**, 383 (1871)).

Die für endliche Gruppen allgemein definierten Begriffe können für die symmetrische Gruppe verwendet werden. Ihre *Einheit*, die mit 1 bezeichnet wird, ist die *identische* Permutation $\begin{pmatrix} 1\,2\,3\ldots n \\ 1\,2\,3\ldots n \end{pmatrix}$, die alle Elemente ungeändert läßt. Die zu

$A = \begin{pmatrix} 1 & 2 & 3 & \dots & n \\ \alpha_1 & \alpha_2 & \alpha_3 & \dots & \alpha_n \end{pmatrix}$ *reziproke* oder *inverse Permutation* lautet $A^{-1} = \begin{pmatrix} \alpha_1 & \alpha_2 & \alpha_3 & \dots & \alpha_n \\ 1 & 2 & 3 & \dots & n \end{pmatrix}$. Die *Ordnung einer Permutation* ist die kleinste positive Zahl a, für die $A^a = 1$ wird.

Die Ordnung irgendeiner Permutation ist das kleinste gemeinsame Vielfache der Anzahl von Symbolen ihrer einzelnen Zykeln. Im besonderen ist die Ordnung einer regulären Permutation gleich der Anzahl von Symbolen irgendeines ihrer Zykeln.

Ist A irgendeine Permutation der Ordnung n und d der größte gemeinsame Teiler der positiven ganzen Zahlen x und n, so hat die Permutation A^x die Ordnung $\frac{n}{d}$; ist A zyklisch, so ist A^x regulär und zerfällt in d Zykeln von je $\frac{n}{d}$ Symbolen. Jede reguläre Permutation ist die Potenz einer zyklischen Permutation.

Auch eine geringere Anzahl als die Gesamtheit aller Permutationen der symmetrischen Gruppe der Ordnung $n!$ kann für sich eine Gruppe bilden. Jede solche Untergruppe der symmetrischen Gruppe heißt eine *Permutationsgruppe.* Die Anzahl der Elemente, auf die sich die Gruppe bezieht, heißt ihr *Grad,* die Anzahl der Permutationen, die sie enthält, ihre *Ordnung. Die Ordnung jeder Permutationsgruppe des Grades n ist ein Divisor von $n!$*

Enthält eine Permutationsgruppe eine Permutation der Klasse k und (abgesehen von der in jeder Gruppe vorhandenen Identität) keine von niedrigerer Klasse, so heißt die *Permutationsgruppe von der Klasse k.* (C. Jordan, *Journ. de math.* (2) **16**, 408 (1871)).

Alle geraden Permutationen von n Symbolen bilden eine Permutationsgruppe der Ordnung $\frac{n!}{2}$, die alternierende Gruppe.

Enthält eine Permutationsgruppe des Grades n die $n-2$ zyklischen Permutationen zweier fester Symbole mit den übrigen, also z. B. $(1, 2, 3), (1, 2, 4), \dots (1, 2, n)$, so ist sie entweder die alternierende oder die symmetrische Gruppe.

Enthält eine Permutationsgruppe des Grades n die $n-1$ Transpositionen eines festen Symbols mit den übrigen, also z. B. $(1, 2), (1, 3), \dots, (1, n)$, so ist sie die symmetrische Gruppe.

Die abstrakte Gruppe, die durch die $n-1$ unabhängigen Elemente $A_i (i = 1, 2, \ldots, n-1)$ mit den Gleichungen

$$A_1^2 = A_2^2 = \cdots = A_{n-1}^2 = 1,$$

$$A_i A_j = A_j A_i (i = 1, 2, \ldots, n-3;\ j = i+2,\ i+3, \ldots, n-1),$$

$$A_j A_{j+1} A_j = A_{j+1} A_j A_{j+1} (j = 1, 2, \ldots, n-2)$$

definiert wird, ist mit der symmetrischen Gruppe holoedrisch isomorph. Analoge Definition einer abstrakten Gruppe der Ordnung $\frac{n!}{2}$, die mit der alternierenden Gruppe des Grades n holoedrisch isomorph ist, durch $n-2$ unabhängige Elemente. (E. H. Moore, *Proc. Lond. M. S.* 28, 357 (1897), Dickson, *Linear groups*, S. 287.)

Zwei Permutationen A und B von n Elementen heißen *konjugiert*, *ähnlich* oder *gleichberechtigt* — schärfer ausgedrückt: konjugiert in bezug auf die symmetrische Gruppe —, falls irgendeine Permutation C des nämlichen Grades existiert, daß $A = C^{-1}BC$ ist. *Zwei Permutationen sind dann und nur dann in bezug auf die symmetrische Gruppe konjugiert, falls sie gleichviele Zykeln mit gleichvielen Elementen besitzen.* Die Permutation $C^{-1}BC$ wird dadurch erhalten, daß man die Permutation C in den Zykeln der Permutation B ausführt. Teilt man die symmetrische Gruppe in k Klassen konjugierter Elemente, so umfaßt die ϱ^{te} Klasse $\frac{n!}{1^\alpha \cdot \alpha!\ 2^\beta \cdot \beta!\ 3^\gamma \cdot \gamma! \ldots}$ Permutationen, die aus α Zykeln mit einem Element, β Zykeln mit zwei Elementen, γ Zykeln mit drei Elementen, ... bestehen. Die Anzahl k der Klassen konjugierter Elemente der symmetrischen Gruppe des Grades n ist gleich der Zahl der ganzzahligen positiven Lösungen $\alpha, \beta, \gamma, \ldots$ der Gleichung $n = \alpha + 2\beta + 3\gamma + \cdots$, wobei $\alpha, \beta, \gamma, \ldots$ auch Null sein können. (Cauchy, *Oeuvres* (1) 9, 289).

Zwei Permutationen A und B, die einer Permutationsgruppe $\mathfrak{G}$ angehören, heißen *in bezug auf* $\mathfrak{G}$ *konjugiert*, *ähnlich* oder *gleichberechtigt*, falls in $\mathfrak{G}$ eine Permutation C existiert, daß $A = C^{-1}BC$ wird. Z. B. zwei Permutationen, die aus lauter Zykeln verschiedener ungerader Ordnungen bestehen und in bezug auf die symmetrische Gruppe ähnlich sind, brauchen es nicht in bezug auf die alternierende Gruppe zu sein. (Frobenius, *Sitzungsb. d. Berl. Akad.* (1901), 303.)

Zwei Permutationsgruppen $\mathfrak{H}_1$ und $\mathfrak{H}_2$ gleichen Grades

heißen *ähnlich, konjugiert*, auch *gleichberechtigt*, falls eine Permutation R des nämlichen Grades existiert, daß $\mathfrak{H}_2 = R\mathfrak{H}_1 R^{-1}$ wird. (Betrachtet man eine Permutationsgruppe als eine Gruppe linearer homogener Substitutionen, so kann man den Begriff der Ähnlichkeit zweier Permutationsgruppen weitergehend mittelst einer überführenden linearen homogenen Substitution definieren, ohne daß eine überführende Permutation zu existieren braucht. W. Burnside, *Proc. Lond. M. S.* **34**, 159 (1902).)

Ist $\mathfrak{H}$ eine Untergruppe der Permutationsgruppe $\mathfrak{G}$ und sind alle Permutationen G von $\mathfrak{G}$ mit $\mathfrak{H}$ vertauschbar, so ist $\mathfrak{H}$ eine *invariante Untergruppe* von $\mathfrak{G}$.

Die symmetrische Gruppe hat die alternierende Gruppe zur invarianten Untergruppe des Index 2.

Die alternierende Gruppe von mehr als vier Symbolen ist einfach (C. Jordan, *C. R.* **60**, 773 (1865), Kronecker, *Monatsb. d. Berl. Akad.* (1879), 208). Die alternierende Gruppe $\mathfrak{G}_{12}$ des Grades 4 besitzt eine invariante Untergruppe $\mathfrak{G}_4$ des Index 3; sie ist zugleich invariante Untergruppe der symmetrischen Gruppe $\mathfrak{G}_{24}$ und besteht aus der Identität und drei Paaren von Transpositionen (1, 2) (3, 4), (1, 3) (2, 4), (1, 4) (2, 3), von denen jede mit der Identität eine invariante Untergruppe von $\mathfrak{G}_4$ bildet. Die symmetrische Gruppe von 4, 3 und 2 Symbolen ist eine auflösbare Gruppe und besitzt die Indexreihe 2, 3, 2, 2 bezw. 2, 3 bezw. 2.

§ 7. Transitive, intransitive, primitive und imprimitive Permutationsgruppen. Reguläre Gruppen. Gruppen vom Primzahlgrad. Metazyklische Gruppe. Modulargruppe.

Eine Permutationsgruppe heißt *transitiv,* wenn ihre Permutationen irgendein Symbol in jedes überzuführen gestatten; eine transitive Permutationsgruppe besitzt Permutationen, die *jedes* Symbol in *jedes* überführen. Eine Permutationsgruppe n^{ten} Grades ist dann und nur dann transitiv, wenn sie n Permutationen der Form: $\begin{pmatrix}1 \dots\\ 1 \dots\end{pmatrix}, \begin{pmatrix}1 \dots\\ 2 \dots\end{pmatrix}, \begin{pmatrix}1 \dots\\ 3 \dots\end{pmatrix}, \dots \begin{pmatrix}1 \dots\\ n \dots\end{pmatrix}$ enthält.

Eine Permutationsgruppe heißt *intransitiv,* wenn die in ihr enthaltenen Permutationen wenigstens ein Symbol nicht in jedes beliebige überführen. Bei jeder intransitiven Gruppe $\mathfrak{G}$ können die Symbole so in m Teile *„Systeme der Intransitivität"*:

$$a_1, a_2, \dots; b_1, b_2, \dots; c_1, c_2, \dots; \dots; m_1, m_2, \dots,$$

zerlegt werden, daß die Permutationen der Gruppe die Symbole jedes Teiles transitiv untereinander vertauschen und nicht in diejenigen eines anderen überführen. Jede Permutation einer intransitiven Gruppe ist das Produkt $P_1 P_2 \dots P_m$ von Permutationen $P_i (i = 1, 2, \dots, m)$, von denen jede nur die Symbole aus einem System enthält. Die Permutationen P_i, die nur die Symbole des i^{ten} Systems der Intransitivität enthalten, bilden eine transitive Permutationsgruppe $\mathfrak{P}_i$. Die Permutationsgruppen $\mathfrak{P}_1, \mathfrak{P}_2, \dots, \mathfrak{P}_m$ heißen die *transitiven Komponenten* (Konstituenten) von $\mathfrak{G}$. Über intransitive Permutationsgruppen vgl. Bolza, *Am. J. math.* **11**, 195 (1889), Burnside, *Theory of groups*, S. 159.

Lassen sich die Symbole einer transitiven Permutationsgruppe des Grades n in m Systeme von je $\frac{n}{m}$ Symbolen verteilen, so daß die Permutationen der Gruppe die Symbole eines Systems entweder nur durcheinander oder durch sämtliche eines anderen Systems ersetzen, so heißt die Gruppe *imprimitiv*. Die m Systeme werden *Systeme der Imprimitivität* genannt. Eine transitive Gruppe, deren Symbole sich nicht so einteilen lassen, heißt *primitiv*.

Die Einteilung der Permutationsgruppen in die drei Klassen: intransitive, imprimitive und primitive geht auf Ruffini zurück.

Ist $\mathfrak{G}$ eine transitive Permutationsgruppe des Grades n und bilden die Permutationen von $\mathfrak{G}$, die ein bestimmtes Symbol ungeändert lassen, eine Gruppe $\mathfrak{H}$, so ist $\mathfrak{G}$ imprimitiv, wenn eine von $\mathfrak{G}$ und $\mathfrak{H}$ verschiedene Permutationsgruppe $\mathfrak{M}$ existiert, die $\mathfrak{H}$ enthält und in $\mathfrak{G}$ enthalten ist. Existiert kein solches $\mathfrak{M}$, so ist $\mathfrak{H}$ primitiv. (Dyck, *Math. Ann.* **22**, 94 (1883), Frobenius, *Sitzungsb. d. Berl. Akad.* (1895), 179.)

Jede invariante Untergruppe einer primitiven Permutationsgruppe ist transitiv. (Satz von Jordan, *Traité*, 41.) *Die Ordnung einer auflösbaren primitiven Permutationsgruppe ist eine Primzahlpotenz.* (Galois, *Œuvres*, p. 11.)

Besitzt eine imprimitive Permutationsgruppe eine invariante intransitive Untergruppe, so sind ihre Systeme der Intransitivität Systeme der Imprimitivität der gegebenen Gruppe. Besitzt eine imprimitive Gruppe $\mathfrak{G}$ m Systeme der Imprimitivität, so werden diese durch die Permutationen von $\mathfrak{G}$ nach einer transitiven Gruppe $\mathfrak{G}_1$ des Grades m permutiert, die mit der Faktorgruppe $\mathfrak{G}/\mathfrak{J}$ holoedrisch isomorph ist. Die invariante Untergruppe $\mathfrak{J}$ von $\mathfrak{G}$, die der Identität von $\mathfrak{G}_1$ entspricht, ist intransitiv und läßt die

m Systeme ungeändert. Die Gruppe $\mathfrak{H}$ kann auch unter Umständen die Identität sein.

Bei einer imprimitiven Permutationsgruppe können sich die Symbole verschiedenartig in Systeme der Imprimitivität zerlegen lassen (vgl. über imprimitive Gruppen die mit zahlreichen Literaturangaben versehene Arbeit von Kuhn, *Am. J. math.* **26**, 45 (1904)).

Die transitiven Permutationsgruppen werden in *einfach* und *mehrfach transitive* unterschieden. (Mathieu, *Journ. de math.* (2) **5**, 13 (1860).) Eine Permutationsgruppe heißt *k-fach transitiv*, wenn ihre Permutationen irgend k fest gewählte Symbole in k beliebige andere überzuführen gestatten. Eine k-fach transitive Gruppe besitzt eine Permutation, die k *beliebig ausgewählte* Symbole in k *beliebige* andere überführt. *Eine mehrfach transitive Gruppe ist stets primitiv.*

Die Ordnung einer k-fach $(k \geqq 1)$ *transitiven Permutationsgruppe vom Grade n ist gleich* $n(n-1)(n-2)\ldots(n-k+1)m$, *worin m die Ordnung einer Untergruppe bedeutet, deren Permutationen k beliebig ausgewählte Symbole ungeändert lassen.*

Ist die Gruppe $(n-k)^{ten}$ *Grades, die von allen denjenigen Permutationen einer wenigstens k-fach transitiven Permutationsgruppe* n^{ten} *Grades gebildet wird, die k bestimmte Symbole ungeändert lassen, noch* μ*-fach transitiv, so ist die Gruppe* n^{ten} *Grades* $(\mu + k)$*-fach transitiv.* (Frobenius, *Journ. f. Math.* **101**, 290 (1887).)

Eine Permutationsgruppe des Grades n, die nicht die alternierende Gruppe ihres Grades enthält, kann nicht mehr als $\left(\frac{n}{3}+1\right)$*-fach transitiv sein*[1]*, die alternierende Gruppe ist* $(n-2)$*-fach, die symmetrische Gruppe n-fach transitiv.*

Enthält eine k-fach transitive Permutationsgruppe nicht die alternierende Gruppe ihres Grades, so ist ihre Klasse (vgl. S. 207) $> 2k-3$. *Besitzt also eine k-fach transitive Permutationsgruppe, abgesehen von der Identität, Permutationen, die weniger als* $2k-2$ *Symbole vertauschen, so ist sie die alternierende oder symmetrische Gruppe.*

Enthält eine k-fach transitive Permutationsgruppe des Grades n nicht die alternierende Gruppe ihres Grades, so ist ihre Klasse $> \frac{1}{4}n-1$, *wenn* $k>1$, $> \frac{1}{3}n-1$, *wenn* $k>2$ *und* $\geqq \frac{1}{2}n-1$,

1) Für die fünffach transitive Mathieusche Permutationsgruppe des Grades 12 (vgl. S. 202) wird dieses Maximum erreicht.

wenn $k > 3$. (Bochert, *Math. Ann.* **40**, 179 (1892); engere Grenzen: *Math. Ann.* **49**, 133 (1897), C. Jordan, *Journ. de math.* (5) **1**, 35 (1895), Maillet, *Mém. prés. par divers savants à l'acad. des sciences* **32**, (1902)).

In jeder transitiven Permutationsgruppe des Grades n gibt es wenigstens $n - 1$ *Permutationen, die alle Symbole vertauschen.* Eine transitive Permutationsgruppe, bei der die Ordnung gleich dem Grade n ist, heißt *regulär. Eine solche enthält, abgesehen von der Identität, nur Permutationen, die alle Symbole vertauschen; sie läßt sich auch als eine transitive Permutationsgruppe des Grades und der Klasse n charakterisieren. Jede reguläre Gruppe, deren Grad eine zusammengesetzte Zahl ist, ist imprimitiv.*

Zu jeder regulären Permutationsgruppe des Grades n existiert eine mit ihr holoedrisch isomorphe derselben Symbole, sie umfaßt alle Permutationen und auch keine anderen als diejenigen, die mit jeder Permutation der ersten Gruppe vertauschbar sind. Der Durchschnitt der zwei Gruppen ist die Zentrale jeder von beiden.

Bei einer transitiven Permutationsgruppe $\mathfrak{G}$ *des Grades n und der Klasse* $n - 1$ *bilden die* $n - 1$ *Permutationen, die alle Symbole umsetzen, zusammen mit der identischen Permutation eine charakteristische Untergruppe. Eine derartige Gruppe* $\mathfrak{G}$ *kann nur dann primitiv sein, wenn n eine Potenz einer Primzahl und die in* $\mathfrak{G}$ *enthaltene Untergruppe* $\mathfrak{N}$ *der Ordnung n eine elementare ist. Unter diesen Bedingungen ist* $\mathfrak{G}$ *stets dann und nur dann primitiv, wenn* $\mathfrak{N}$ *eine minimale invariante Untergruppe von* $\mathfrak{G}$ *ist.* (Frobenius, *Sitzungsb. d. Berl. Akad.* (1901), 1226, (1902), 458.)

Eine Permutationsgruppe des Grades $r + k$, *die nicht die alternierende Gruppe ihres Grades enthält, kann, wenn r irgendeine Primzahl ist, für* $k > 2$ *nicht mehr als k-fach transitiv sein.* (C. Jordan, *Bull. de la soc. math.* **1**, 41, G. A. Miller, *Bull. Am. M. S.* **4**, 142 (1898).)

Für *transitive Permutationsgruppen vom Primzahlgrade p* gelten folgende Sätze:

Eine transitive Permutationsgruppe vom Primzahlgrad ist stets primitiv.

Ist der Grad einer transitiven Permutationsgruppe $\mathfrak{G}$ *eine Primzahl p, so ist ihre Ordnung* $= pq(\lambda p + 1)$, *wo q ein Teiler von* $p - 1$ *ist und* $\lambda p + 1$ *die Anzahl der verschiedenen in* $\mathfrak{G}$ *enthaltenen Untergruppen* $\mathfrak{P}$ *der Ordnung p bedeutet; diese Untergruppen sind sämtlich konjugiert. Alle mit einer beliebigen derartigen Gruppe* $\mathfrak{P}$ *vertauschbaren Elemente von* $\mathfrak{G}$ *bilden eine*

Untergruppe der Ordnung pq von $\mathfrak{G}$. (Mathieu, *Journ. de math.* (2) **6**, 304 (1861).)

Enthält eine transitive Permutationsgruppe $\mathfrak{G}$ vom Primzahlgrade p irgendeine invariante Untergruppe $\mathfrak{J}$, so ist $\mathfrak{G}/\mathfrak{J}$ eine zyklische Gruppe. (G. A. Miller, *Bull. Am. M. S.* **4**, 141 (1898).)

Die Permutationen

$$\begin{pmatrix} 0 & 1 & 2 & \dots & p-1 \\ \alpha & \alpha+\beta & \alpha+2\beta & \dots & \alpha+(p-1)\beta \end{pmatrix},$$

wobei $\alpha = 0, 1, 2, \dots, p-1$, $\beta = 1, 2, \dots, p-1$ und die im Nenner der Permutation stehenden Zahlen immer durch ihre kleinsten Reste mod p zu ersetzen sind, bilden eine *Permutationsgruppe des Grades p und der Ordnung $p(p-1)$. Sie ist zweifach transitiv und auflösbar. Jede auflösbare transitive Permutationsgruppe vom Primzahlgrad p ist entweder diese Gruppe oder eine ihrer Untergruppen.* (Galois, *Œuvres*, p. 47.) Die fragliche Gruppe der Ordnung $p(p-1)$ heißt nach Kronecker (*Monatsb. d. Berl. Akad.* (1879), 217) *metazyklisch*, bei H. Weber, *Algebra* **1**, 666 die *volle lineare Gruppe*. Man stellt ihre Permutationen *„analytisch"* dar durch die Relation $|z, \alpha z + \beta|$, welche, wenn z die Werte $0, 1, 2, \dots, p-1$ durchläuft, die Verknüpfung der im Zähler der Permutation stehenden Symbole mit den unter ihnen befindlichen des Nenners angibt. Die Gruppe wird erzeugt durch die zwei Permutationen $|z, z+1|$ und $|z, gz|$, wobei g eine primitive Wurzel der Primzahl p ist. Verallgemeinerung dieser Gruppe zu einer zweifach transitiven Permutationsgruppe des Grades p^m und der Ordnung $p^m(p^m-1)$ mit p^m konjugierten zyklischen Untergruppen der Ordnung $p^m - 1$, indem für die Primzahl p die p^m Zahlen des $GF[p^m]$ treten. Mathieu, *Journ. de math.* (2) **5**, 39 (1860), (2) **6**, 262 (1861) (eine weitere Verallgemeinerung), Burnside, *Theory of groups*, S. 155, de Séguier, *Journ. de math.* (5) **8**, 263 (1902).

Jede transitive Permutationsgruppe vom Primzahlgrade p ist entweder in der metazyklischen Gruppe enthalten oder wenigstens zweifach transitiv. W. Burnside, *Proc. Lond. M. S.* **33**, 163 (1901), *Quart. J.* **37**, 215 (1906), J. Schur, *Math. Ver.* **17**, 171 (1908).

Die metazyklische Gruppe und ihre Untergruppen, deren es für jeden Divisor q von $p-1$ eine gibt, sind die einzigen transitiven Permutationsgruppen vom Primzahlgrad p mit einer einzigen Untergruppe $(\lambda = 0)$ der Ordnung p. Die Zahl λ muß

immer Null sein, wenn $q = 1$ ist oder, wenn für $q = 2$, die Primzahl p von der Form $4t + 3$ ist.

Für $\lambda = 1$ gilt der Satz:

Es gibt nur vier transitive Permutationsgruppen, deren Grad eine Primzahl p ist und die $p + 1$ Untergruppen der Ordnung p enthalten, die alternierende und die symmetrische Gruppe des Grades 5, *deren Ordnungen gleich* 60 *und* 120 *sind, und die beiden einfachen Gruppen der Grade* 7 *und* 11, *deren Ordnungen gleich* 168 *und* 660 *sind.* (Sylow, *Videnskabsselskabets Skrifter* I (1897), Frobenius, *Sitzungsb. d. Berl. Akad.* (1902), 352.)

Die Permutationen in $p + 1$ Symbolen, die mit $0, 1, 2, \ldots, p - 1, \infty$ bezeichnet seien,

$$\begin{pmatrix} 0 & 1 & 2 & 3 & \ldots & p-1 & \infty \\ \frac{\beta}{\delta} & \frac{\alpha+\beta}{\gamma+\delta} & \frac{2\alpha+\beta}{2\gamma+\delta} & \frac{3\alpha+\beta}{3\gamma+\delta} & \cdots & \frac{(p-1)\alpha+\beta}{(p-1)\gamma+\delta} & \frac{\alpha}{\gamma} \end{pmatrix}$$

bilden, wenn p eine ungerade Primzahl, $\alpha, \beta, \gamma, \delta$ beliebige Zahlen aus der Reihe $0, 1, 2, \ldots, p - 1$ sind, deren Determinante $\alpha\delta - \beta\gamma \not\equiv 0$ ist, und die im Nenner der Permutation stehenden Zahlen mod p genommen werden, eine Gruppe; hierbei ist unter $\frac{a}{b}$ in üblicher Weise die Zahl x zu verstehen, die durch $bx \equiv a \pmod{p}$ bestimmt wird, und für $b = 0$, $a \not\equiv 0$ ist $\frac{a}{b} = \infty$ zu setzen.

Die definierte Permutationsgruppe des Grades $p + 1$ hat die Ordnung $p(p^2 - 1)$ und ist dreifach transitiv. Ihre Permutationen werden „analytisch" dargestellt durch die Relation $\left| z, \frac{\alpha z + \beta}{\gamma z + \delta} \right|$, $\alpha\delta - \beta\gamma \not\equiv 0$; diese gibt, wenn z die Werte $0, 1, 2, \ldots, p - 1, \infty$ durchläuft, die Verknüpfung der im Zähler der Permutation stehenden Symbole mit den unter ihnen befindlichen des Nenners an. *Die Gruppe der Ordnung $p(p^2 - 1)$ hat eine invariante, zweifach transitive Untergruppe des Grades $p + 1$ und der Ordnung $\frac{1}{2}p(p^2 - 1)$.* Ihre Permutationen werden definiert durch $\left| z, \frac{\alpha z + \beta}{\gamma z + \delta} \right|$ mit der Bedingung $\alpha\delta - \beta\gamma \equiv 1 \pmod{p}$. *Die Gruppe der Ordnung $\frac{1}{2}p(p^2 - 1)$ ist für $p > 3$ einfach.* Die letztere Gruppe ergibt sich als Galoissche Gruppe bei den Transformationsgleichungen der elliptischen Funktionen (Modulargleichungen) und heißt daher die *Modulargruppe.* Sie ist bereits von Galois,

Œuvres, p. 28, behandelt. Die Bestimmung der Untergruppen der Modulargruppe bei Gierster, *Math. Ann.* **18**, 319 (1881), vgl. Klein-Fricke, *Modularfunktionen* **1**, 411, Weber, *Algebra* **3**, 284.

Ist p eine Primzahl, so gibt es nicht mehr als eine transitive Gruppe des Grades $p+1$ und der Ordnung $\frac{1}{2}p(p^2-1)$; nur für $p=7$ gibt es zwei solche. Es gibt nicht mehr als eine transitive Gruppe des Grades $p+1$ und der Ordnung $p(p^2-1)$. (Frobenius, *Sitzungsb. d. Berl. Akad.* (1902), 353, 359.)

Die durch die Permutationen $\left| z, \frac{\alpha z+\beta}{\gamma z+\delta} \right|$ $(\alpha\delta-\beta\gamma \neq 0)$ definierte Gruppe des Grades $p+1$ läßt sich zu einer *dreifach transitiven Permutationsgruppe L des Grades p^m+1 und der Ordnung $p^m(p^{2m}-1)$* verallgemeinern, indem man sich statt der Primzahl p des Galoisschen Feldes $GF[p^m]$ bedient. Als die p^m+1 zu vertauschenden Symbole hat man die mit 0, 1, 2, ..., p^m-1 numerierten Zahlen des $GF[p^m]$ unter Zufügung des Symbols ∞ zu verstehen, $\alpha, \beta, \gamma, \delta$ durchlaufen die Zahlen des $GF[p^m]$, unter $\frac{a}{b}$ sei die durch $b\xi = a$ definierte Zahl ξ des $GF[p^m]$ verstanden. Ist $p=2$, so ist die Gruppe L mit der Gruppe L_0 der Permutationen $\left| z \frac{\alpha z+\beta}{\gamma z+\delta} \right|$, $\alpha\delta-\beta\gamma=+1$ identisch. Für $p>2$ hat die Gruppe L die Gruppe L_0 des Grades p^m+1 und der Ordnung $\frac{p^m(p^{2m}-1)}{2}$, die durch die Permutationen $\left| z \frac{\alpha z+\beta}{\gamma z+\delta} \right|$, $\alpha\delta-\beta\gamma=+1$ definiert wird und zweifach transitiv ist, zur invarianten Untergruppe. Die Permutationsgruppe L_0 des Grades p^m+1 und der Ordnung $\frac{p^m(p^{2m}-1)}{2}$ bezw. $2^m(2^{2m}-1)$ heißt die *verallgemeinerte endliche Modulargruppe.* (Mathieu, *Journ. de math.* (2) **5**, 39 (1860), E. H. Moore, *Math. papers of the Chicago Congress* (1893), publ. 1896, S. 208, *The decennial publications of the university of Chicago,* Vol. IX (1903), W. Burnside, *Proc. Lond. M. S.* **25** 113 (1894), Wiman, *Stockholm Akad. Bihang* **25** (1899), J. Schur, *Journ. f. Math.* **132**, 113 (1907), Dickson, *Linear groups,* S. 260, H. Weber, *Algebra* **2**, 310.) *Die verallgemeinerte Modulargruppe ist, abgesehen von den zwei Fällen $p^m=2$ und $p^m=3$, eine einfache Gruppe. Sie hat stets Untergruppen des Index p^m+1, aber nur für $p^m=2, 3, 5, 7, 3^2, 11$ solche von niederem Index, nämlich vom Index 2, 3, 5, 7, 6, 11.* (Für $m=1$ bereits

bei Galois, *Œuvres,* p. 29.) Außer für $p^m = 5, 7, 9, 11$ ist die verallgemeinerte Modulargruppe mit keiner Permutationsgruppe niedrigeren Grades als mit einer des Grades $p^m + 1$ holoedrisch isomorph. In den Ausnahmefällen, wo ihre Ordnung 60, 168, 360 und 660 ist, kann sie auch als Permutationsgruppe in 5, 7, 6 und 11 Symbolen dargestellt werden und ist dann 3-, 2-, 4- und 2-fach transitiv.

Die Gruppe L der Ordnung $p^m(p^{2m} - 1)$ hat die oben S. 213 behandelte verallgemeinerte metazyklische Gruppe $|z, \alpha z + \beta|$ der Ordnung $p^m(p^m - 1)$ zur Untergruppe. (G. A. Miller, *Quart. J.* **34**, 232 (1903).)

§ 8. Permutationsgruppen höchster Ordnung bei gegebenem Grad. Die zu den Permutationsgruppen zugehörigen Funktionen. Die symmetrischen und alternierenden Funktionen.

Die Ordnung einer Permutationsgruppe n^{ten} Grades ist stets ein Teiler von $n!$, jedoch kann nicht jeder Teiler von $n!$ die Ordnung einer Permutationsgruppe n^{ten} Grades sein.

Die Aufgabe, Untergruppen der symmetrischen Gruppe von möglichst kleinem Index zu bestimmen, wird als *Bertrandsches Problem* (Bertrand, *J. éc. polyt. Cah.* **30**, 123 (1845)) bezeichnet. *Der Index einer intransitiven Untergruppe der symmetrischen Gruppe n^{ten} Grades ist gleich oder größer als n und nur dann gleich n, wenn die Permutationsgruppe ein Symbol fest läßt und die übrigen $n-1$ Symbole auf alle $(n-1)!$ Arten permutiert. Der Index einer imprimitiven Untergruppe der symmetrischen Gruppe von n Symbolen ist immer größer als n, abgesehen von $n = 4$. Für $n = 4$ existieren drei ähnliche imprimitive Gruppen 8ten Grades, also vom Index 3, von denen eine durch die zwei Permutationen* $(12)(3)(4), (13)(24)$ *erzeugt wird. Ausgenommen für $n = 6$ gibt es keine primitive Permutationsgruppe in n Symbolen, deren Ordnung $\geqq (n-1)!$ ist. Für $n = 6$ gibt es primitive Permutationsgruppen des Index 6, also von der Ordnung* 120; *eine solche wird durch die auf S. 214 besprochene der Primzahl $p = 5$ entsprechende Gruppe $(p+1)^{\text{ten}}$ Grades der Ordnung $p(p^2-1)$ geliefert.* (Vgl. J. A. Serret, *Journ. de math.* **15**, 1 (1850), *Algèbre supérieure* **2**, 287, C. Jordan, *Traité,* 67, H. Weber, *Algebra* **2**, 154.)

Eine Aufzählung aller Permutationsgruppen einschließlich der intransitiven, deren Ermittlung sich stets auf transitive niedrigeren Grades zurückführen läßt, ist bis zum Grade $n = 11$ geleistet; bei den imprimitiven Permutationsgruppen ist man

bis zum Grade $n = 15$, bei den primitiven[1]) bis $n = 18$ gegangen. Literatur in der nicht abgeschlossenen Arbeit von E. N. Martin, *Am. J. math.* **23**, 285 (1901). Zur Ergänzung: Miller and Ling, *Quart. J.* **32**, 342 (1901), Kuhn, *Am. J. math.* **26**, 45 (1904). Vgl. auch die Zusammenstellung bei Easton, *The constructive development of group-theory*, p. 77.

Hat man eine rationale Funktion $\varphi(x_1, x_2, \ldots, x_n)$ der n Variablen $x_1, x_2, \ldots, x_n$ und führt auf sie sämtliche $n!$ Permutationen der n Zahlen $1, 2, \ldots, n$ aus, so bildet die Gesamtheit aller Permutationen, bei denen sich die Funktion φ nicht ändert, eine Permutationsgruppe $\mathfrak{G}$ n^{ten} Grades; sie heißt die *Gruppe der Funktion. Umgekehrt gehören zu jeder Permutationsgruppe* $\mathfrak{G}$ n^{ten} *Grades in den Zahlen* $1, 2, \ldots, n$ *eine und sogar unendlich viele rationale Funktionen der* n *Variablen* $x_1, x_2, \ldots, x_n$, *die nur bei den Permutationen aus* $\mathfrak{G}$ *und keinen anderen ungeändert bleiben.*

Eine rationale Funktion $\varphi(x_1, x_2, \ldots, x_n)$, die bei den Permutationen der symmetrischen Gruppe $\mathfrak{S}$ ϱ verschiedene Werte annimmt, heißt ϱ-*wertig.* Ist $\mathfrak{G}$ die Gruppe der Funktion φ, hat $\mathfrak{G}$ die Ordnung g und geht φ durch die Permutationen von $\mathfrak{S}$ in die ϱ verschiedenen Werte $\varphi_1 = \varphi, \varphi_2, \varphi_3, \ldots, \varphi_\varrho$ über, so ist $\varrho = \frac{n!}{g}$, *also gleich dem Index der Untergruppe* $\mathfrak{G}$ *von* $\mathfrak{S}$. (Lagrange.) Die ϱ Funktionen $\varphi_1, \varphi_2, \ldots, \varphi_\varrho$ heißen ein *System* (in bezug auf $\mathfrak{S}$) *konjugierter Funktionen.*

Jede Permutation aus $\mathfrak{S}$ permutiert die ϱ Funktionen untereinander. Geht φ_a $(a = 1, 2, \ldots, \varrho)$ durch eine Permutation A von $\mathfrak{S}$ in φ_b über und ist $\mathfrak{G}_1$ die Gruppe der Funktion φ_a, so geht φ_a nur durch die Permutationen der Nebengruppe $\mathfrak{G}_1 A$ in φ_b über und $A^{-1}\mathfrak{G}_1 A$ ist die zu φ_b gehörige Gruppe. *Zu konjugierten Funktionen gehören innerhalb* $\mathfrak{S}$ *ähnliche Permutationsgruppen.*

Da, abgesehen von $n = 4$, die symmetrische Gruppe außer der alternierenden Gruppe des Index 2 keine invariante Untergruppe besitzt, haben für $\varrho > 2$, abgesehen von $n = 4$, die zu einem System ϱ konjugierter Funktionen zugehörigen Permutations-

---

1) Primitive Permutationsgruppen ungerader Ordnung, deren Grad < 243 ist, existieren, abgesehen von Untergruppen der metazyklischen Gruppe, nur 10 der Ordnungen: 25.3, 27.13, 27.39, 81.5, 121.3, 121.15, 125.31, 125.93, 169.7, 169.21. Der erste Faktor gibt hierbei den Grad der betr. Permutationsgruppe an. H. L. Rietz, *Am. J. math.* **26**, 30 (1904).

gruppen keine andere gemeinschaftliche Permutation als die Einheit.

Abgesehen von $n = 4$, ist eine weniger als n-wertige Funktion von n Variablen 1- oder 2-wertig. (Folge des oben besprochenen Bertrandschen Problems.)

Für $n = 4$ gibt es dreiwertige Funktionen, nämlich das System konjugierter Funktionen:

$$\varphi_1 = \{(x_1 + x_2) - (x_3 + x_4)\}^2, \quad \varphi_2 = \{(x_1 + x_3) - (x_2 + x_4)\}^2,$$
$$\varphi_3 = \{(x_1 + x_4) - (x_2 + x_3)\}^2.$$

Die drei zu diesen Funktionen zugehörigen ähnlichen $\mathfrak{G}_8$ haben die in der symmetrischen Gruppe $\mathfrak{G}_{24}$ vierten Grades vorhandene invariante $\mathfrak{G}_4$ (vgl. S. 209) gemeinsam. Hierauf beruht Lagranges Auflösung der Gleichung vierten Grades (Lagrange, *Œuvres* **3**, 266, vgl. das Kapitel über Algebra § 8).

Die einwertigen rationalen Funktionen von n Größen heißen *symmetrisch. Jede rationale symmetrische Funktion ist der Quotient zweier ganzer rationaler symmetrischer Funktionen.*

Unter der symmetrischen Funktion $\sum x_1^{\nu_1} x_2^{\nu_2} \ldots x_n^{\nu_n}$ ist der Gliederkomplex zu verstehen, der gefunden wird, wenn man $x_1^{\nu_1} x_2^{\nu_2} \ldots x_n^{\nu_n}$ allen Permutationen der symmetrischen Gruppe unterwirft und die Summe aller dieser Ausdrücke bildet, dabei aber mehrfach auftretende fortläßt. Anders ausgedrückt:

$$\sum x_1^{\nu_1} x_2^{\nu_2} \ldots x_n^{\nu_n}$$

ist die Summe des Systems der mit $x_1^{\nu_1} x_2^{\nu_2} \ldots x_n^{\nu_n}$ konjugierten Funktionen.

Bei $\sum x_1^{\nu_1} x_2^{\nu_2} \ldots x_n^{\nu_n}$ kann man das Glied, dessen Exponenten den Ungleichungen $\nu_1 \geqq \nu_2 \geqq \nu_3 \ldots \geqq \nu_n$ genügen, zum Anfangsglied wählen. Ist dies geschehen, so bezeichnet man $\sum x_1^{\nu_1} x_2^{\nu_2} \ldots x_n^{\nu_n}$ nur durch Angabe der Exponenten $(\nu_1, \nu_2, \ldots, \nu_n)$ (G. Cramer, *Introduction à l'analyse des lignes courbes*, Genf 1750, p. 666, Vandermonde, *Résolution des équations* (1771), deutsche Ausg., Berlin 1888, S. 9). Man nennt $(\nu_1, \nu_2, \ldots, \nu_n)$ eine *typische, einfache* oder *eintypige symmetrische Funktion.* Von zwei typischen symmetrischen Funktionen $(\nu_1, \nu_2, \ldots, \nu_n)$ und $(\nu_1', \nu_2', \ldots \nu_n')$ heißt die *erste von höherer Ordnung als die zweite*, wenn in der Zahlenreihe $\nu_1 - \nu_1'$, $\nu_2 - \nu_2'$, $\ldots$, $\nu_n - \nu_n'$ die erste nicht verschwindende Differenz positiv ist, also entweder $\nu_1 > \nu_1'$ oder $\nu_1 = \nu_1'$, $\nu_2 > \nu_2'$ oder $\nu_1 = \nu_1'$, $\nu_2 = \nu_2'$, $\nu_3 > \nu_3'$, usw. (Gauß, *Zweiter Beweis des Fundamentalsatzes*

der Algebra (1815), *Ges. Werke* **3**, 36, *Ostwalds Klassiker der exakten Wiss.* Nr. 14, herausg. von Netto, S. 40.)

Die typischen symmetrischen Funktionen niedrigster Ordnung sind:

$$S_1 = (1, 0, 0, \ldots, 0) = \sum x_1 = x_1 + x_2 + \cdots + x_n,$$

$$S_2 = (1, 1, 0, 0, \ldots, 0) = \sum x_1 x_2 = x_1 x_2 + x_1 x_3 + \cdots + x_1 x_n + x_2 x_3 + \cdots + x_{n-1} x_n,$$

$$S_3 = (1, 1, 1, 0, \ldots, 0) = \sum x_1 x_2 x_3 = x_1 x_2 x_3 + x_1 x_2 x_4 + \cdots + x_{n-2} x_{n-1} x_n,$$

$$\vdots$$

$$S_n = (1, 1, 1, \ldots, 1) = x_1 x_2 x_3 \ldots x_n.$$

Die n Funktionen $S_1, S_2, \ldots, S_n$ heißen die *elementaren symmetrischen* oder die *symmetrischen Grundfunktionen.*

Jede symmetrische ganze Funktion $G(x_1, x_2, \ldots, x_n)$ läßt sich als eine Summe typischer symmetrischer Funktionen der Form $(\nu_1, \nu_2, \ldots, \nu_n)$ mit von den Größen $x_1, x_2, \ldots, x_n$ unabhängigen Zahlenkoeffizienten $M_{\nu_1 \nu_2 \ldots \nu_n}$ darstellen,

$$G = \sum M_{\nu_1 \nu_2 \ldots \nu_n} (\nu_1, \nu_2, \ldots, \nu_n).$$

Es ist (Waring, *Meditationes algebraicae* (1782), 3. ed., p. 13)

$$(\nu_1, \nu_2, \ldots, \nu_n) = S_1^{\nu_1 - \nu_2} S_2^{\nu_2 - \nu_3} \ldots S_n^{\nu_n} + G_1;$$

hierbei bedeutet G_1 eine symmetrische Funktion, die sich nur aus typischen symmetrischen Funktionen *niedrigerer Ordnung* als $(\nu_1, \nu_2, \ldots, \nu_n)$ zusammensetzt. Hieraus ergibt sich der Hauptsatz (Cramer, Vandermonde, Waring, a. a. O., Gauß, *Ges. Werke* **3**, 36, *Ostwalds Klassiker der exakten Wiss.* Nr. 14, S. 40, Cauchy, *Exercices de math.* **4** (1829), *Œuvres* (2) **9**, 132, Kronecker, *Ges. Werke* **1**, 323, **2**, 290):

Jede ganze symmetrische Funktion der Größen $x_1, x_2, \ldots, x_n$ mit ganzzahligen Koeffizienten ist auf eine und auch nur auf eine Weise als ganze ganzzahlige Funktion der symmetrischen Grundfunktionen $S_1, S_2, \ldots, S_n$ darstellbar.

Im besonderen drücken sich die ganzen *Potenzsummen*

$$s_\nu = (\nu, 0, 0, \ldots, 0) = x_1^\nu + x_2^\nu + \cdots + x_n^\nu$$

in der Form

$$s_\nu = \nu \sum (-1)^{\lambda_2 + \lambda_4 + \lambda_6 + \cdots} \frac{(\lambda_1 + \lambda_2 + \cdots + \lambda_n - 1)!}{\lambda_1! \lambda_2! \ldots \lambda_n!} S_1^{\lambda_1} S_2^{\lambda_2} \ldots S_n^{\lambda_n}$$

aus; hierbei ist das Summenzeichen rechter Hand über alle po-

sitiven ganzen Zahlen einschließlich 0 *zu erstrecken, die der Gleichung* $\lambda_1 + 2\lambda_2 + 3\lambda_3 + \cdots + n\lambda_n = \nu$ *genügen* (Waringsche *Formeln, Miscellanea analytica* (1762), vgl. Saalschütz, *Bibliotheca math.* (3) **9**, 65 (1908)) Als Umkehrung ergibt sich (Waring, *Miscellanea analytica*):

$$S_\nu = \sum \frac{(-1)^{\nu+\lambda_1+\lambda_2+\cdots+\lambda_\nu}}{1^{\lambda_1} \cdot 2^{\lambda_2} \cdot 3^{\lambda_3} \ldots \nu^{\lambda_\nu}} \frac{1}{\lambda_1! \lambda_2! \ldots \lambda_\nu!} \cdot s_1^{\lambda_1} s_2^{\lambda_2} \ldots s_\nu^{\lambda_\nu};$$

hierbei ist $\lambda_1 + 2\lambda_2 + \cdots + \nu\lambda_\nu = \nu$.

Die Darstellung der typischen symmetrischen Funktionen durch die ganzen Potenzsummen wird durch die ebenfalls sogenannten *Waringschen Formeln* gegeben (Waring, *Miscellanea analytica* (1762)). Sie lauten

$$(\nu_1, \nu_2, 0, 0, \ldots, 0) = s_{\nu_1} \cdot s_{\nu_2} - s_{\nu_1+\nu_2} \quad (\nu_1 \neq \nu_2),$$
$$(\nu_1, \nu_1, 0, 0, \ldots, 0) = \tfrac{1}{2} s_{\nu_1}^2 - \tfrac{1}{2} s_{2\nu_1}, \quad \text{usw.}$$

Diese Formeln hat Faà di Bruno, *Einleitung in die Theorie der binären Formen,* deutsch von Walter, Leipzig 1881, S. 8 in die Form folgender symbolischer Determinante gebracht:

$$(\nu_1, \nu_2, \ldots, \nu_k, 0, 0, 0 \ldots, 0) = \begin{vmatrix} s_{\nu_1} & s^{\nu_1} & s^{\nu_1} & \ldots & s^{\nu_1} \\ s^{\nu_2} & s_{\nu_2} & s^{\nu_2} & \ldots & s^{\nu_2} \\ \vdots & \vdots & \vdots & & \vdots \\ s^{\nu_k} & s^{\nu_k} & s^{\nu_k} & \ldots & s_{\nu_k} \end{vmatrix};$$

nach der Entwicklung sind die Exponenten der s durch Indices zu ersetzen. Sind unter den Größen $\nu_1, \nu_2, \ldots, \nu_k$ gleiche vorhanden, so ist für je l gleiche die rechte Seite durch $l!$ zu dividieren.

Über symmetrische Funktionen vergleiche auch das Kapitel über Algebra § 5.

Eine Funktion der n Variablen $x_1, x_2, \ldots, x_n$ heißt *alternierend,* wenn sie bei allen Permutationen der n Zahlen 1, 2, $\ldots, n$ nur ihr Vorzeichen ändert.

Jede alternierende Funktion ist von der Form $G\Delta$, *wobei* G *eine symmetrische Funktion und* Δ *das Differenzenprodukt der* n *Variablen* $x_1, x_2, \ldots, x_n$ *bedeutet* (vgl. S. 68).

Jede zweiwertige Funktion ist von der Form $G_1 + G_2 A$, *wobei* G_1 *und* G_2 *symmetrische Funktionen und* A *eine alternierende Funktion bedeutet. Umgekehrt ist jede Funktion der angegebenen Form zweiwertig.*

Ist $\varphi(x_1, x_2, \ldots, x_n)$ *eine* ϱ*-wertige Funktion, so sind* φ *und die mit* φ *konjugierten Funktionen Wurzeln einer Gleichung* ϱ^{ten} *Grades:*

$$z^\varrho + G_1 z^{\varrho-1} + G_2 z^{\varrho-2} + \cdots + G_\varrho = 0,$$

wobei $G_1, G_2, \ldots, G_\varrho$ *symmetrische Funktionen von* $x_1, x_2, \ldots, x_n$ *bedeuten.*

Ist $\mathfrak{G}$ *die Gruppe der rationalen Funktion* $\varphi(x_1, x_2, \ldots, x_n)$ *der* n *Variablen* $x_1, x_2, \ldots, x_n$ *und bleibt die ebenfalls rationale Funktion* $\psi(x_1, x_2, \ldots, x_n)$ *bei allen Permutationen von* $\mathfrak{G}$ *ungeändert, so ist* ψ *eine rationale Funktion von* φ *mit Koeffizienten, die symmetrische Funktionen von* $x_1, x_2, \ldots, x_n$ *sind* (Lagrange, *Œuvres* **3**, 374 (1770)). *Ist* $\mathfrak{H}_h$ *die zu* ψ *gehörige Gruppe*[1]) *und* $\sigma = \frac{h}{g}$ *der Index der Untergruppe* $\mathfrak{G}_g$ *von* $\mathfrak{H}_h$, *so genügt* φ *einer Gleichung* σ^{ten} *Grades*

$$z^\sigma + G_1 z^{\sigma-1} + G_2 z^{\sigma-2} + \cdots + G_\sigma = 0,$$

wobei $G_1, G_2, \ldots, G_\sigma$ *rationale Funktionen von* ψ *und den symmetrischen Funktionen von* $x_1, x_2, \ldots, x_n$ *bedeuten. Die Wurzeln der Gleichung* σ^{ten} *Grades sind die mit* φ *„in bezug auf die Gruppe* $\mathfrak{H}$ *konjugierten Funktionen", d. h. die verschiedenen aus* $\varphi(x_1, x_2, \ldots, x_n)$ *hervorgehenden Größen, wenn man* φ *den Permutationen der Gruppe* $\mathfrak{H}$ *unterwirft.*

Sieht man $x_1, x_2, \ldots x_n$ als die Wurzeln einer „allgemeinen" Gleichung n^{ten} Grades an, d. h. einer solchen, deren Koeffizienten willkürliche Größen sind und deren Galoissche Gruppe demnach die symmetrische Permutationsgruppe n^{ten} Grades ist, so erhält man die obigen Sätze als Spezialfälle derjenigen Sätze, die in der Galoisschen Theorie (vgl. *Algebra*) für irgendwelche Gleichungen n^{ten} Grades mit beliebigen speziellen Koeffizienten hergeleitet werden.

Über die zu einer Permutationsgruppe zugehörigen Funktionen sei noch verwiesen auf den Artikel *„Rationale Funktionen der Wurzeln; symmetrische und Affektfunktionen"* von Vahlen in der *Encyklopädie der math. Wiss.* **1**, 449.

---

1) Die Gruppe $\mathfrak{H}$ enthält $\mathfrak{G}$ als Untergruppe oder ist mit $\mathfrak{G}$ identisch. Um die σ verschiedenen mit φ in bezug auf die Gruppe $\mathfrak{H}$ konjugierten Funktionen zu finden, genügt es, die Funktion $\varphi(x_1, x_2, \ldots, x_n)$ je einer Substitution aus den σ Nebengruppen zu unterwerfen, die man erhält, wenn man die Gruppe $\mathfrak{H}$ nach dem Modul $\mathfrak{G}$ zerlegt.

§ 9. Lineare homogene Substitutionsgruppen. Reduzibilität. Endliche und unendliche Gruppen. Invarianten endlicher Gruppen.

Ein System $\mathfrak{G}$ von Matrices gleichen Grades n von nicht verschwindenden Determinanten bildet eine Gruppe, wenn es von der Vollständigkeit ist, daß das Produkt irgend zweier sowie die reziproke zu jeder in $\mathfrak{G}$ enthaltenen Matrix der Gesamtheit $\mathfrak{G}$ angehört. Da zu jeder Matrix eine lineare homogene Substitution und umgekehrt (vgl. S. 82) gehört, so spricht man von einer *Gruppe linearer homogener Substitutionen.* (Eine weitergehende Definition, die nicht das Nichtverschwinden der Determinanten der Matrices fordert, bei Frobenius u. J. Schur, *Sitzungsb. der Berl. Akad.* (1906), 209.) Die Zahl n heißt *der Grad der Substitutionsgruppe.*

Umfaßt $\mathfrak{G}$ die Gesamtheit *aller* linearen homogenen Substitutionen von nicht verschwindenden Determinanten, so heißt $\mathfrak{G}$ die *allgemeine lineare homogene Gruppe*, von der jede lineare homogene Substitutionsgruppe Untergruppe ist.

Die Gruppen linearer homogener Substitutionen lassen sich in *reduzible* und *irreduzible* einteilen. Eine Gruppe $\mathfrak{G}$ linearer homogener Substitutionen n^{ten} Grades heißt *reduzibel*, wenn irgendeine Matrix P von nicht verschwindender Determinante existiert, daß die Matrix jeder Substitution der mit $\mathfrak{G}$ ähnlichen Gruppe $\mathfrak{A} = P\mathfrak{G}P^{-1}$ die Gestalt

$$\begin{matrix} R & 0 \\ S & T \end{matrix}$$

hat, wo die den verschiedenen Substitutionen von $\mathfrak{A}$ zugehörigen Matrices R sämtlich von gleichem, aber von niedrigerem als n^{tem} Grade sind. In dem Fall schreibt man auch: die Gruppe $\mathfrak{A}$ hat die Form

$$\begin{matrix} \mathfrak{A}_{11} & 0 \\ \mathfrak{A}_{21} & \mathfrak{A}_{22}, \end{matrix}$$

wobei $\mathfrak{A}_{11}$ die Gesamtheit aller Matrices R, $\mathfrak{A}_{21}$ aller Matrices S und $\mathfrak{A}_{22}$ aller Matrices T repräsentiert. $\mathfrak{A}_{11}$ und $\mathfrak{A}_{22}$ definieren mit $\mathfrak{G}$ und $\mathfrak{A}$ homomorphe Gruppen. Jede nicht reduzible Gruppe heißt *irreduzibel.*

Jede Gruppe $\mathfrak{G}$ linearer homogener Substitutionen läßt sich mittelst einer Matrix P von nicht verschwindender Determinante in eine ähnliche Gruppe $\mathfrak{A} = P\mathfrak{G}P^{-1}$ transformieren, daß die Matrix jeder Substitution von $\mathfrak{A}$ die Form hat:

$$\begin{matrix} \mathfrak{a}_{11} & 0 & 0 & \dots & 0 \\ \mathfrak{a}_{21} & \mathfrak{a}_{22} & 0 & \dots & 0 \\ \mathfrak{a}_{31} & \mathfrak{a}_{32} & \mathfrak{a}_{33} & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ \mathfrak{a}_{i1} & \mathfrak{a}_{i2} & \mathfrak{a}_{i3} & \dots & \mathfrak{a}_{ii} \end{matrix}$$

und sämtliche in der Diagonale stehenden Matrices $\mathfrak{a}_{11}, \mathfrak{a}_{22}, \dots, \mathfrak{a}_{ii}$ *irreduzible Gruppen* definieren. Die Gruppen $\mathfrak{a}_{11}, \mathfrak{a}_{22}, \dots, \mathfrak{a}_{ii}$ heißen die *irreduziblen Bestandteile, die sich bei der Darstellung von* $\mathfrak{G}$ *in der Form* $\mathfrak{A}$ *ergeben.*

Wie auch immer eine Gruppe $\mathfrak{G}$ *linearer homogener Substitutionen unter Hervorhebung ihrer irreduziblen Bestandteile in eine ähnliche Gruppe transformiert wird, so lassen sich die irreduziblen Bestandteile, die sich bei irgendeiner Darstellung ergeben, den irreduziblen Bestandteilen, die sich bei irgendeiner anderen Darstellung ergeben, eineindeutig so zuordnen, daß zwei zugeordnete irreduzible Teilgruppen gleichviele Variablen haben und ähnliche Gruppen sind.* (A. Loewy, *Trans. Am. M. S.* 4, 44 (1903)). Dieser Satz läßt sich auf folgende Weise verallgemeinern (Frobenius und J. Schur, *Sitzungsber. d. Berl. Akad.* (1906), 215): *Zwei holoedrisch isomorphe Gruppen linearer homogener Substitutionen besitzen dann und nur dann die nämlichen irreduziblen Bestandteile, wenn die Spuren*[1]) *je zweier einander entsprechender Substitutionen der beiden isomorphen Gruppen übereinstimmen.*

Eine Gruppe $\mathfrak{G}$ linearer homogener Substitutionen heißt *vollständig reduzibel*, wenn man wenigstens eine Matrix P von nicht verschwindender Determinante finden kann, daß die zu $\mathfrak{G}$ ähnliche Gruppe $\mathfrak{A} = P\mathfrak{G}P^{-1}$ die besondere Form $\{\mathfrak{a}_{11}, \mathfrak{a}_{22}, \dots, \mathfrak{a}_{ii}\}$:

$$\begin{matrix} \mathfrak{a}_{11} & 0 & 0 & \dots & 0 \\ 0 & \mathfrak{a}_{22} & 0 & \dots & 0 \\ 0 & 0 & \mathfrak{a}_{33} & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & \mathfrak{a}_{ii} \end{matrix}$$

1) Unter der Spur einer linearen homogenen Substitution versteht man die Summe der Wurzeln der charakteristischen Funktion der betreffenden Substitution.

hat, also abgesehen von der Diagonale, die lauter irreduzible Gruppen enthält, nur Nullmatrices stehen. *Auch die irreduziblen Gruppen* ($i = 1$) *werden zu den vollständig reduziblen Gruppen gerechnet.*

Nicht jede Gruppe linearer homogener Substitutionen ist vollständig reduzibel. Jede Gruppe $\mathfrak{G}$ linearer homogener Substitutionen läßt sich durch eine Substitution A von nicht verschwindender Determinante in eine ähnliche Gruppe $\mathfrak{G}^* = A\mathfrak{G}A^{-1}$ überführen, bei der die Matrix jeder Substitution die Form:

$$\begin{matrix} \mathfrak{G}_{11}^* & 0 & 0 & \dots & 0 \\ \mathfrak{G}_{21}^* & \mathfrak{G}_{22}^* & 0 & \dots & 0 \\ \mathfrak{G}_{31}^* & \mathfrak{G}_{32}^* & \mathfrak{G}_{33}^* & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ \mathfrak{G}_{\mu 1}^* & \mathfrak{G}_{\mu 2}^* & \mathfrak{G}_{\mu 3}^* & \dots & \mathfrak{G}_{\mu\mu}^* \end{matrix}$$

hat und $\mathfrak{G}_{11}^*, \mathfrak{G}_{22}^*, \dots, \mathfrak{G}_{\mu\mu}^*$ *aufeinanderfolgende größte vollständig reduzible Gruppen* bedeuten. *Betrachtet man ähnliche Gruppen linearer homogener Substitutionen als nicht verschieden, so sind die Gruppen* $\mathfrak{G}_{11}^*, \mathfrak{G}_{22}^*, \dots, \mathfrak{G}_{\mu\mu}^*$ *ihrer Reihenfolge nach als erste, zweite usw. bis* μ^{te} *vollständig reduzible, zu* $\mathfrak{G}$ *gehörige Gruppen eindeutig bestimmt und von der Wahl der überführenden Matrix* A *unabhängig. Zu jeder Gruppe linearer homogener Substitutionen gehört daher eine homomorphe vollständig reduzible Gruppe* $\{\mathfrak{G}_{11}^*, \mathfrak{G}_{22}^*, \dots, \mathfrak{G}_{\mu\mu}^*\}$ *desselben Grades. Sieht man ähnliche Gruppen als nicht verschieden an, so ist diese ebenso wie ihre irreduziblen Bestandteile*[1]) *eindeutig bestimmt.* (A. Loewy, *Trans. Am. M. S.* **6**, 504 (1905), Stickelberger, ebenda **7**, 509 (1906), J. Schur, erscheint ebenda **10** (1909), A. Loewy, *Math. Ann.* **64**, 267 (1907).) Die Gruppe $\mathfrak{G}$ ist dann und nur dann vollständig reduzibel, wenn $\mu = 1$ ist.

Eine Gruppe linearer homogener Substitutionen in n Variablen ist dann und nur dann irreduzibel, wenn aus ihr ein System von n^2 *Substitutionen* $A_j (j = 1, 2, \dots, n^2)$ *mit Koeffizienten* $a_{ik}^{(j)}$ $(i, k = 1, 2, \dots, n;\ j = 1, 2, \dots, n^2)$ *ausgewählt werden kann, daß die Determinante:*

---

1) Die irreduziblen Bestandteile der Gruppe $\mathfrak{G}$ sind die Gesamtheit der irreduziblen Bestandteile der vollständig reduziblen Gruppen $\mathfrak{G}_{11}^*, \mathfrak{G}_{22}^*, \dots, \mathfrak{G}_{\mu\mu}^*$.

$$\begin{vmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \dots & a_{nn}^{(1)} \\ a_{11}^{(2)} & a_{12}^{(2)} & \dots & a_{nn}^{(2)} \\ \vdots & & & \\ a_{11}^{(n^2)} & a_{12}^{(n^2)} & \dots & a_{nn}^{(n^2)} \end{vmatrix} \neq 0$$

ist. Anders ausgedrückt: Eine Gruppe n^{ten} Grades ist dann und nur dann irreduzibel, wenn unter ihren Matrices n^2 linear unabhängige (vgl. S. 86) *existieren.* (W. Burnside, *Proc. Lond. M. S.* (2) **3**, 430 (1905), Frobenius und J. Schur, *Sitzungsb. d. Berl. Akad.* (1906), 209, Taber, *Math. Ann.* **64**, 357 (1907).)

Die *irreduziblen,* auch *transitiv* genannten Gruppen linearer homogener Substitutionen lassen sich in *primitive* und *imprimitive* einteilen. Eine irreduzible Gruppe linearer homogener Substitutionen heißt *imprimitiv*, wenn sie in eine derartige ähnliche Gruppe transformierbar ist, daß die Variablen in m Systeme von je $\frac{n}{m}$ Variablen zerfallen und alle Variablen eines Systems entweder nur in lineare homogene Funktionen der Variablen des gleichen oder eines anderen Systems übergehen. (Blichfeldt, *Trans. Am. M. S.* **4**, 388 (1903), **6**, 230 (1905)).

Zu den reduziblen Gruppen gehören auch die kommutativen Gruppen linearer homogener Substitutionen; für sie gilt folgendes Theorem: *Jede kommutative Gruppe linearer homogener Substitutionen, die durch q Elemente erzeugt wird, ist reduzibel, und jeder ihrer irreduziblen Bestandteile ist eine Gruppe in einer einzigen Variablen.* Anders ausgedrückt: *Die Matrices aller Substitutionen einer solchen Gruppe lassen sich durch eine und dieselbe Matrix in dreieckige Form transformieren, so daß rechter Hand von der Diagonale ausschließlich Nullen auftreten.* (Frobenius, *Sitzungsb. d. Berl. Akad.* (1896), 601, J. Schur, ebenda (1902), 120, H. Weber, *Algebra* **2**, 176, Dickson, *Quart. J.* **40**, 167 (1909).) *Die Anzahl der linear unabhängigen Matrices einer kommutativen Gruppe linearer homogener Substitutionen n^{ten} Grades ist höchstens*

$$\left[\frac{n^2}{4}\right] + 1,$$

wobei $\left[\frac{n^2}{4}\right]$ die größte in $\frac{n^2}{4}$ enthaltene ganze Zahl ist. (J. Schur, *Journ. f. Math.* **130**, 66 (1905).)

Die Gruppen linearer homogener Substitutionen werden in *endliche* und *unendliche* unterschieden. Eine Gruppe $\mathfrak{G}$ linearer

homogener Substitutionen heißt von der *endlichen Ordnung g*, wenn sie nur die endliche Anzahl g linearer homogener Substitutionen umfaßt.

Bei jeder endlichen Gruppe $\mathfrak{G}$ linearer homogener Substitutionen ist jede einzelne Substitution von endlicher Ordnung. Umgekehrt: *Weiß man, daß alle Substitutionen einer Gruppe linearer homogener Substitutionen von niedrigerer als q^{ter} Ordnung sind, wobei q irgendeine positive Zahl ist, so ist die Gruppe endlich. Eine unendliche Gruppe linearer homogener Substitutionen muß wenigstens eine Substitution unendlich hoher Ordnung enthalten.* (A. Loewy, *Math. Ann.* **53**, 225 (1900), **64**, 264 (1907), W. Burnside, *Proc. Lond. M. S.* (2) **3**, 435 (1905).)

Eine lineare homogene Substitution ist dann und nur dann von endlicher Ordnung, wenn ihre charakteristische Funktion bloß für Einheitswurzeln verschwindet und sämtliche Elementarteilerexponenten gleich 1 *sind. Die Matrix einer jeden linearen homogenen Substitution n^{ten} Grades von endlicher Ordnung s ist daher mit der zerlegbaren Matrix* $\{\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n\}$ *ähnlich* (vgl. S. 92), *wobei* $\varepsilon_i^s = 1$ $(i = 1, 2, \ldots, n)$ *ist.* (Frobenius, *Journ. f. Math.* **84**, 16 (1878), *Sitzungsb. d. Berl. Akad.* (1896), 607, C. Jordan, *Journ. f. Math.* **84**, 112 (1878), vgl. auch H. Weber, *Algebra* **2**, 186.)

Jede endliche Gruppe linearer homogener Substitutionen führt eine definite Hermitesche Form (d. h. eine Hermitesche Form (vgl. S. 128), die nur verschwindet, wenn sämtliche Variablen gleich Null gesetzt werden) *in sich über.* (A. Loewy. *C. R.* **123**, 168 (1896), E. H. Moore, *Math. Ann.* **50**, 213 (1898).) Die verwandten Untersuchungen von L. Fuchs, *Sitzungsb. d. Berl. Akad.* (1896), 753 sind fehlerhaft; vgl. die Bemerkungen von Schlesinger zu diesem Aufsatz in *Fuchs' Ges. Werken* **3** (erscheint nächstens).

Jede endliche Gruppe linearer homogener Substitutionen ist vollständig reduzibel. (H. Maschke, *Math. Ann.* **52**, 363 (1899), W. Burnside, *Acta math.* **28**, 377 (1904), J. Schur, *Sitzungsb. d. Berl. Akad.* (1905), 414, A. Loewy, *Trans. Am. M. S.* **6**, 509 (1905).)

Bisher ist keine endliche Gruppe linearer homogener Substitutionen g^{ter} Ordnung bekannt, die nicht einer solchen ähnlich ist, deren Koeffizienten ausnahmslos dem Zahlkörper der g^{ten} Einheitswurzeln angehören. (Über die erzielten Resultate vgl. Maschke, *Math. Ann.* **50**, 492 (1898), W. Burnside,

Proc. Lond. M. S. (2) **3**, 239 (1905), J. Schur, *Sitzungsb. d. Berl. Akad.* (1906), 164.)

Jede endliche Gruppe linearer homogener Substitutionen von Primzahlpotenzordnung ist einer Gruppe mit Substitutionen der Form (sog. Monomialgruppe) $x_i = a_i x_{\lambda_i}$ $(i = 1, 2, \ldots, n)$ *ähnlich, wobei* $a_1, a_2, \ldots, a_n$ *Konstante sind und* $\lambda_1, \lambda_2, \ldots, \lambda_n$ *bis auf die Reihenfolge die Zahlen* $1, 2, \ldots, n$ *bedeuten.* (Blichfeldt, *Trans. Am. M. S.* **5**, 313 (1904).)

Satz von C. Jordan (*Journ. f. Math.* **84**, 91 (1878)):

Jede endliche Gruppe $\mathfrak{G}$ *linearer homogener Substitutionen in n Variablen besitzt eine invariante kommutative Untergruppe* $\mathfrak{J}$ *von der Beschaffenheit, daß der Quotient* $\lambda = \frac{g}{i}$ *der Ordnungen von* $\mathfrak{G}$ *und* $\mathfrak{J}$ *kleiner als eine nur durch n allein bestimmte Zahl ist.* Neuer Beweis mit Bestimmung einer oberen Grenze für λ bei Blichfeldt, *Trans. Am. M. S.* **4**, 387 (1903), **5**, 310 (1904), **6**, 230 (1905). Die Ordnung $g = \lambda i$ kann beliebig große Werte annehmen (vgl. J. Schur, *Sitzungsb. d. Berl. Akad.* (1905), 77).

Eine ganze homogene Funktion $\Phi(x_1 x_2 \ldots x_n)$ der n Variablen $x_1, x_2, \ldots, x_n$ (Form) heißt eine *Invariante einer Gruppe* $\mathfrak{G}$ *linearer homogener Substitutionen*, wenn sie bis auf einen von den Variablen $x_1, x_2, \ldots, x_n$ unabhängigen Faktor ungeändert bleibt, falls man die x_i allen Substitutionen der Gruppe unterwirft. Man unterscheidet zwischen *relativen* und *absoluten Invarianten einer Gruppe* $\mathfrak{G}$. *Die Faktoren, mit denen sich eine relative Invariante einer Gruppe* $\mathfrak{G}$ *linearer homogener Substitutionen von der endlichen Ordnung g multiplizieren kann, sind ausnahmslos* g^{te} *Einheitswurzeln. Sind alle Faktoren, mit denen sich eine relative Invariante* $\Phi(x_1, x_2, \ldots, x_n)$ *einer* $\mathfrak{G}_g$ *multipliziert,* e^{te} *Einheitswurzeln und ist wenigstens ein Faktor auch keine niedrigere als* e^{te} *Einheitswurzel, so heißt* Φ *eine Invariante der Gruppe* $\mathfrak{G}$ *vom Index e; die Gruppe* $\mathfrak{G}$ *enthält alsdann eine invariante Untergruppe* $\mathfrak{J}$, *und der Index von* $\mathfrak{J}$ *in bezug auf* $\mathfrak{G}$ *ist gleich e.*

Ein System von Formen heißt endlich, wenn sich jede Form des Systems als ganze rationale Funktion einer endlichen Anzahl von Formen, die ebenfalls dem System angehören, darstellen läßt. Diese endliche Anzahl von Formen, durch die sich jede Form des Systems in ganzer rationaler Weise ausdrücken läßt, heißt ein *volles Formensystem.*

Das System aller Invarianten jeder endlichen Gruppe linearer homogener Substitutionen ist endlich. Das System aller absoluten

Invarianten jeder endlichen Gruppe linearer homogener Substitutionen ist endlich. (Hilberts Satz von der Endlichkeit des Invariantensystems einer endlichen Gruppe, Hilbert, *Math. Ann.* **36**, 473 (1890), H. Weber, *Algebra* **2**, 218.)

Unter den Invarianten jeder endlichen Gruppe linearer homogener Substitutionen, die relative Invarianten besitzt, kann man m Invarianten $\Psi_1, \Psi_2, \ldots, \Psi_m$ derartig auswählen, daß jede Invariante der Gruppe von der Form: $\Phi_1\Psi_1 + \Phi_2\Psi_2 + \cdots \Phi_m\Psi_m$ wird, wobei $\Phi_1, \Phi_2, \ldots, \Phi_m$ absolute Invarianten sind.

Ist $\mathfrak{G}$ die symmetrische Gruppe $\mathfrak{S}_{n!}$ von n Symbolen, so bilden die symmetrischen Grundfunktionen $S_1, S_2, \ldots, S_n$ ein volles System absoluter Invarianten, das Differenzenprodukt Δ ist eine relative Invariante vom Index 2 und bildet mit $S_1, S_2, \ldots, S_n$ ein volles System aller Invarianten von $\mathfrak{S}_{n!}$ (Hieraus ergeben sich auch die Sätze auf S. 219 u. 220.)

Als Aufgabe der Algebra kann angesehen werden: Man soll die Unbekannten $x_1, x_2, \ldots, x_n$ finden, wenn die Invarianten $S_1, S_2, \ldots, S_n$ gegeben sind. Es handelt sich um die Bestimmung der Wurzeln der Gleichung:

$$z^n - S_1 z^{n-1} + S_2 z^{n-2} - S_3 z^{n-3} \cdots (-1)^n S_n = 0 .$$

Klein (*Math. Ann.* **15**, 253 (1879), *The Evanston Colloquium, Lectures on mathematics*, New York 1894, S. 72) hat dieses Problem dahin erweitert: *Für unbekannte $x_1, x_2, \ldots, x_n$ seien die Werte eines vollen Systems absoluter Invarianten einer endlichen linearen homogenen Substitutionsgruppe $\mathfrak{G}$ in Übereinstimmung mit den zwischen ihnen bestehenden Relationen*[1]) *gegeben: man soll $x_1, x_2, \ldots, x_n$ bestimmen. Die Bestimmung der Variablen $x_1, x_2, \ldots, x_n$ hängt von einer Normalgleichung ab, deren Galoissche Gruppe mit $\mathfrak{G}$ holoedrisch isomorph ist.* (Vgl. H. Weber, *Algebra* **2**, 228.)

§ 10. Darstellung einer endlichen abstrakten Gruppe als Permutationsgruppe und als Gruppe linearer homogener Substitutionen.

Jede Permutationsgruppe, die mit einer gegebenen endlichen abstrakten Gruppe holoedrisch oder meroedrisch isomorph ist, bezeichnet man als eine *Darstellung der abstrakten Gruppe*

1) Zwischen $n+1$ Invarianten einer $\mathfrak{G}$ n^{ten} Grades muß stets eine rationale Gleichung bestehen, wie sich durch Elimination der n Variablen x_i $(i = 1, 2, \ldots, n)$ aus den $n+1$ homogenen ganzen Funktionen ergibt.

als Permutationsgruppe. Sind $G_1, G_2, \ldots, G_g$ *die Elemente einer endlichen abstrakten Gruppe* $\mathfrak{G}$ *und ordnet man dem Element* G_i *entweder die Permutation*

$$\begin{pmatrix} G_1 & G_2 & \ldots & G_g \\ G_1 G_i & G_2 G_i & \ldots & G_g G_i \end{pmatrix}$$

oder

$$\begin{pmatrix} G_1 & G_2 & \ldots & G_g \\ G_i^{-1} G_1 & G_i^{-1} G_2 & \ldots & G_i^{-1} G_g \end{pmatrix}$$

zu, so hat man zwei mit $\mathfrak{G}$ *holoedrisch isomorphe reguläre Permutationsgruppen. Jede Permutation der einen Gruppe ist mit jeder der anderen vertauschbar.* (Cayley, *Am. J. math.* **1**, 52 (1878), *Coll. math. papers* **10**, 403, Frobenius u. Stickelberger, *Journ. f. Math.* **86**, 230 (1879), Dyck, *Math. Ann.* **22**, 84 (1883).)

Hat die Gruppe $\mathfrak{G}$ *der Ordnung* g *die Gruppe* $\mathfrak{H}$ *der Ordnung* h *zur Untergruppe und ist* $\mathfrak{J}$ *die größte in* $\mathfrak{H}$ *enthaltene invariante Untergruppe von* $\mathfrak{G}$, *so ist* $\mathfrak{G}/\mathfrak{J}$ *als holoedrisch isomorphe transitive Permutationsgruppe* $\mathfrak{h}$ *in* $\frac{g}{h}$ *Symbolen darstellbar. Ist die größte in* $\mathfrak{H}$ *enthaltene invariante Untergruppe von* $\mathfrak{G}$ *die Einheit, so ist* $\mathfrak{h}$ *eine holoedrisch isomorphe Darstellung von* $\mathfrak{G}$ *selbst. In diesem Fall ist* $\mathfrak{h}$ *eine primitive Permutationsgruppe, wenn es keine Gruppe* $\mathfrak{A}$ *gibt, die in* $\mathfrak{G}$ *enthalten ist und* $\mathfrak{H}$ *enthält; existiert eine derartige Gruppe* $\mathfrak{A}$, *so ist* $\mathfrak{h}$ *imprimitiv.* (Dyck, *Math. Ann.* **22**, 94 (1883), Frobenius, *Sitzungsb. d. Berl. Akad.* (1895), 179.)

Jede einfache Gruppe ist als holoedrisch isomorphe primitive Permutationsgruppe darstellbar. Eine solche Darstellung erhält man nach dem vorigen Satz, indem man $\mathfrak{H}$ als Maximaluntergruppe von $\mathfrak{G}$ wählt.

Über die Darstellbarkeit einer endlichen abstrakten Gruppe als Permutationsgruppe vgl.: W. Burnside, *Proc. Lond. M. S.* **34**, 159 (1902), G. A. Miller, *Giornale di mat.* **38**, 63 (1900).

Sind $G_1, G_2, \ldots, G_g$ die g Elemente einer endlichen abstrakten Gruppe $\mathfrak{G}$, so heißen g Matrices (lineare homogene Substitutionen) gleichen Grades $(G_1), (G_2), \ldots, (G_g)$, die bei ihrer Komposition den g^2 Relationen: $(R)(S) = (RS)$ genügen, wenn R und S alle Elemente aus $\mathfrak{G}$ durchlaufen, eine *Dar-*

stellung der Gruppe $\mathfrak{G}$ durch lineare homogene Substitutionen; hierbei brauchen die Matrices $(G_1), (G_2), \ldots, (G_g)$ nicht sämtlich untereinander verschieden zu sein. Verschwinden die Determinanten der Matrices $(G_1), (G_2), \ldots, (G_g)$ nicht, so definieren $(G_1), (G_2), \ldots, (G_g)$ eine *mit der abstrakten Gruppe $\mathfrak{G}$ homomorphe Gruppe linearer homogener Substitutionen.*

Die Anzahl verschiedener, d. h. nicht untereinander ähnlicher, mit der Gruppe $\mathfrak{G}$ homomorpher irreduzibler Gruppen linearer homogener Substitutionen ist endlich und gleich der Zahl k von Klassen konjugierter Elemente von $\mathfrak{G}$. Der Grad jeder dieser Substitutionsgruppen ist ein Divisor der Ordnung der Gruppe $\mathfrak{G}$.

Bildet das System von Matrices n^{ten} Grades $(G_1), (G_2), \ldots, (G_g)$ eine Darstellung der abstrakten Gruppe $\mathfrak{G}$ und bedeuten x_{G_i} $(i = 1, 2, \ldots, g)$ g unabhängige Variable, so heißt die Matrix $\sum_R (R) x_R$ $(R = G_1, G_2, \ldots, G_g)$ die *dieser Darstellung entsprechende Gruppenmatrix oder eine zur Gruppe $\mathfrak{G}$ gehörige Matrix.* Die zu einer irreduziblen Gruppe linearer homogener Substitutionen gehörige Gruppenmatrix heißt eine *irreduzible Gruppenmatrix.* Ist X irgendeine Gruppenmatrix und A eine konstante Matrix von nicht verschwindender Determinante, so ist die ähnliche Matrix AXA^{-1} ebenfalls eine Gruppenmatrix. Sieht man ähnliche irreduzible Gruppenmatrices als nicht verschieden an, so gibt es entsprechend den k verschiedenen mit einer abstrakten Gruppe $\mathfrak{G}$ homomorphen irreduziblen Gruppen linearer homogener Substitutionen genau k *irreduzible zur Gruppe $\mathfrak{G}$ gehörige irreduzible Gruppenmatrices.*

Jede Gruppenmatrix X des Grades n und des Ranges r läßt sich in eine ähnliche Gruppenmatrix: $\{U_1, U_2, \ldots, U_m, N_{n-r}\}$ (vgl. die Symbolik auf S. 92) *transformieren, so daß $U_1, U_2, \ldots, U_m$ irreduzible Gruppenmatrices bedeuten und N_{n-r} die Nullmatrix des Grades $n - r$ ist. Wenn man ähnliche irreduzible Gruppenmatrices als nicht verschieden ansieht, entspricht auf diese Weise jeder Gruppenmatrix ein bis auf die Reihenfolge eindeutig bestimmtes System irreduzibler Gruppenmatrices $U_1, U_2, \ldots, U_m$; dieses enthält nur Matrices aus den k zur Gruppe $\mathfrak{G}$ gehörigen irreduziblen Gruppenmatrices und diese eventuell auch mehrfach.*

Die Koeffizienten irgendeiner Gruppenmatrix

$$X = \sum (R) x_R = \| x_{ik} \| \qquad (i, k = 1, 2, \ldots, n)$$

sind lineare homogene Funktionen der Variablen $x_{G_1}, x_{G_2}, \ldots, x_{G_g}$. Ist $y_{G_1}, y_{G_2}, \ldots, y_{G_g}$ ein zweites System g unabhängiger Variablen, und setzt man

$$z_R = \sum_S x_S y_{S^{-1}R} \quad (R, S = G_1, G_2, \ldots, G_g),$$

so besteht die Gleichung $Z = XY$, falls Y und Z die Matrices bedeuten, die aus X hervorgehen, wenn man die Variablen x_R durch y_R bzw. z_R ersetzt.

Jede Permutationsgruppe läßt sich als eine Gruppe linearer homogener Substitutionen auffassen. Infolgedessen gehören zu den am Anfang dieses Paragraphen besprochenen Darstellungen der abstrakten Gruppe $\mathfrak{G}$ durch zwei reguläre Permutationsgruppen auch zwei Gruppenmatrices, nämlich

$$X_1 = \| x_{G_i^{-1}G_k} \| \quad \text{und} \quad X_2 = \| x_{G_i G_k^{-1}} \|.$$

Die Matrix X_2, deren Zeilen und Kolonnen man erhält, indem man füx G_i und G_k der Reihe nach die Elemente $G_1, G_2, \ldots, G_g$ setzt, heißt die *reguläre Gruppenmatrix.* Die zu X_1 transponierte Matrix $X_1' = \| x_{G_k^{-1}G_i} \|$ heißt die *antistrophe Matrix.* Ist Y_1' die aus X_1' hervorgehende Matrix, wenn man die Variablen $x_{G_1}, x_{G_2}, \ldots, x_{G_g}$ durch $y_{G_1}, y_{G_2}, \ldots, y_{G_g}$ ersetzt, so besteht die Gleichung $Y_1' X_2 = X_2 Y_1'$. *Die Elemente $G_1, G_2, \ldots, G_g$ einer endlichen abstrakten Gruppe $\mathfrak{G}$ lassen sich als die g Einheiten eines Systems höherer komplexer Zahlen auffassen.* Daher findet man auch die reguläre Gruppenmatrix X_2 und die Matrix Y_1' durch Spezialisierung der Matrices $\mathfrak{G}$ und $\mathfrak{H}'$ auf S. 99.

Die Determinante der regulären Gruppenmatrix X_2 ist nicht identisch Null; daher transformiert sich X_2 in $\{U_1, U_2, \ldots, U_m\}$, und es tritt keine Nullmatrix auf. *Hat man die reguläre Gruppenmatrix, so sind unter $U_1, U_2, \ldots, U_m$ sämtliche k irreduziblen Gruppenmatrices, und zwar jede so oft, als ihr Grad angibt, enthalten. Die Bestimmung aller Darstellungen einer endlichen abstrakten Gruppe durch ganze lineare homogene Substitutionen ist daher mit der Aufgabe identisch, die reguläre Gruppenmatrix in irreduzible Bestandteile zu zerfällen.*

Bilden $(G_1), (G_2), \ldots . (G_g)$ eine Darstellung der endlichen Gruppe $\mathfrak{G}$ und ist $\chi(R)$ die Spur der Substitution (R), d. h. die Summe der Wurzeln der charakteristischen Funktion der Substitution (R), so heißt das System der g Zahlen $\chi(R)$, das höchstens k untereinander verschiedene enthält, wenn (R) die

Matrices $(G_1), (G_2), \ldots, (G_g)$ durchläuft, der *Gruppencharakter*, der dieser Darstellung oder ihrer zugehörigen Gruppenmatrix entspricht. Ein einer irreduziblen Gruppenmatrix entsprechender Gruppencharakter heißt ein *einfacher Gruppencharakter*. Unter den Zahlen eines einfachen Gruppencharakters befindet sich auch die der Einheit (E) der Gruppe zugeordnete Zahl $\chi(E)$, die den Grad der irreduziblen Gruppe angibt.

Zwei Darstellungen einer endlichen Gruppe $\mathfrak{G}$ durch lineare homogene Substitutionen von nicht verschwindenden Determinanten sind dann und nur dann ähnlich, wenn sie denselben Gruppencharakter besitzen. (Spezialfall des oben S. 223 angegebenen Satzes über isomorphe unendliche Gruppen linearer homogener Substitutionen von Frobenius u. J. Schur, *Sitzungsb. d. Berl. Akad.* (1906), 215.)

Eingehend untersucht ist die Darstellung der symmetrischen und alternierenden Gruppe durch Gruppen linearer homogener Substitutionen (Frobenius, *Sitzungsb. d. Berl. Akad.* (1900), 516, (1903), 328). *Jede Gruppe linearer homogener Substitutionen, die der symmetrischen Permutationsgruppe in n Symbolen homomorph ist, läßt sich durch eine lineare Transformation der Variablen in eine ähnliche Gruppe mit ganzzahligen rationalen Koeffizienten überführen.* (J. Schur, *Sitzungsb. d. Berl. Akad.* (1908), 664.)

Das Problem, die Darstellungen einer endlichen abstrakten Gruppe durch lineare homogene Substitutionen zu finden, ist zuerst von Frobenius, *Sitzungsb. d. Berl. Akad.* (1896), 985 u. 1343, (1897), 994, (1899), 482, (1903), 328, 401 und Molien, *Sitzungsb. d. naturforsch. Gesellsch. zu Dorpat* (1897), 259 durchgeführt worden. Vgl. besonders die elementare Darstellung von J. Schur, *Sitzungsb. d. Berl. Akad.* (1905), 406, ferner W. Burnside, *Acta math.* **28**, 369 (1904), *Proc. Lond. M. S.* (2) **1**, 117 (1904), H. Weber, *Algebra* **2**, 193. In dem Problem, alle mit einer abstrakten Gruppe $\mathfrak{G}$ homomorphen Gruppen linearer homogener Substitutionen zu bestimmen, ist im besonderen das von F. Klein (*The Evanston Colloquium, Lectures on mathematics*, New York 1894, S. 73) sogenannte *Normalproblem* enthalten, das die Bestimmung aller mit einer Gruppe isomorphen linearen homogenen Substitutionsgruppen niedrigsten Grades verlangt.

§ 11. Kollineationsgruppen. Darstellbarkeit einer endlichen abstrakten Gruppe als Kollineationsgruppe. Die verschiedenen Typen endlicher Kollineationsgruppen in einer, zwei und drei Variablen.

Aus zwei linear gebrochenen Substitutionen:

$$A_1 : \xi_i = \frac{a_{i1}\xi_1' + a_{i2}\xi_2' + \cdots + a_{in-1}\xi_{n-1}' + a_{in}}{a_{n1}\xi_1' + a_{n2}\xi_2' + \cdots + a_{nn-1}\xi_{n-1}' + a_{nn}} \quad (i = 1, 2, \ldots, n-1)$$

$$B_1 : \xi_i' = \frac{b_{i1}\xi_1'' + b_{i2}\xi_2'' + \cdots + b_{in-1}\xi_{n-1}'' + b_{in}}{b_{n1}\xi_1'' + b_{n2}\xi_2'' + \cdots + b_{nn-1}\xi_{n-1}'' + b_{nn}} \quad (i = 1, 2, \ldots, n-1)$$

in $n-1$ Variablen entspringt eindeutig eine dritte, ihr *Produkt* $C_1 = A_1 B_1$:

$$C_1 : \xi_i = \frac{c_{i1}\xi_1'' + c_{i2}\xi_2'' + \cdots + c_{in-1}\xi_{n-1}'' + c_{in}}{c_{n1}\xi_1'' + c_{n2}\xi_2'' + \cdots + c_{nn-1}\xi_{n-1}'' + c_{nn}} \quad (i = 1, 2, \ldots, n-1),$$

wobei $c_{ik} = \sum_{s=1}^{s=n} a_{is} b_{sk}$ $(i, k = 1, 2, \ldots, n)$.

Verschwindet die Determinante $|a_{ik}|$ $(i, k = 1, 2, \ldots, n)$ nicht, so existiert zu A_1 eine reziproke Substitution:

$$A_1^{-1} : \xi_i' = \frac{A_{1i}\xi_1 + A_{2i}\xi_2 + \cdots + A_{n-1i}\xi_{n-1} + A_{ni}}{A_{1n}\xi_1 + A_{2n}\xi_2 + \cdots + A_{n-1n}\xi_{n-1} + A_{nn}} \quad (i = 1, 2, \ldots, n-1),$$

wobei die Größen $A_{ik} = \frac{\partial |a_{ik}|}{\partial a_{ik}}$ die Unterdeterminanten von $|a_{ik}|$ $(i, k = 1, 2, \ldots, n)$ bedeuten. Den Grad n der Matrix $\|a_{ik}\|$ bezeichnet man als den *Grad der Substitution* A_1.

Eine Gesamtheit $\mathfrak{G}_1$ linear gebrochener Substitutionen von gleichem Grad und nicht verschwindenden Determinanten bildet eine Gruppe, wenn sie von der Vollständigkeit ist, daß das Produkt irgend zweier sowie die reziproke zu jeder in $\mathfrak{G}_1$ enthaltenen Substitution ebenfalls der Gesamtheit $\mathfrak{G}_1$ angehört. *Eine Gruppe linear gebrochener Substitutionen heißt eine Kollineationsgruppe* oder *projektive Gruppe.* Umfaßt die Gruppe $\mathfrak{G}_1$ die Gesamtheit *aller* linear gebrochenen Substitutionen von gleichem Grad und nicht verschwindenden Determinanten, so heißt $\mathfrak{G}_1$ die *allgemeine projektive Gruppe* (vgl. S. 175) oder die *allgemeinste Kollineationsgruppe*, von der jede Gruppe linear gebrochener Substitutionen Untergruppe ist. Im besonderen gehört zu den Untergruppen der allgemeinen projektiven Gruppe die *allgemeine lineare unhomogene Substitutionsgruppe* in $n-1$

Variablen, auch die *volle lineare Gruppe* in $n-1$ Variablen genannt; sie besteht aus allen Substitutionen:

$$\xi_i = a_{i1}\xi_1' + a_{i2}\xi_2' + \cdots + a_{in-1}\xi_{n-1}' + a_{in} \qquad (i = 1, 2, \ldots, n-1)$$

mit nicht verschwindenden Determinanten $|a_{ik}|$ $(i, k = 1, 2, \ldots, n-1)$.

Ordnet man den Substitutionen A:

$$x_i = a_{i1}x_1' + a_{i2}x_2' + \cdots + a_{in}x_n' \qquad (i = 1, 2, \ldots, n)$$

einer Gruppe $\mathfrak{G}$ linearer homogener Substitutionen die linear gebrochenen Substitutionen A_1:

$$\xi_i = \frac{a_{i1}\xi_1' + a_{i2}\xi_2' + \cdots + a_{in-1}\xi_{n-1}' + a_{in}}{a_{n1}\xi_1' + a_{n2}\xi_2' + \cdots + a_{nn-1}\xi_{n-1}' + a_{nn}} \qquad (i = 1, 2, \ldots, n-1)$$

zu, so erhält man eine mit $\mathfrak{G}$ homomorphe Kollineationsgruppe $\mathfrak{G}_1$, die zu $\mathfrak{G}$ *zugehörige Kollineationsgruppe* $\mathfrak{G}_1$. Ist Γ die Gruppe aller in $\mathfrak{G}$ etwa enthaltenen Ähnlichkeitstransformationen: $x_i = \nu x_i'$ $(i = 1, 2, \ldots, n)$, so sind $\mathfrak{G}/\Gamma$ und $\mathfrak{G}_1$ holoedrisch isomorph.

Nicht jede Kollineationsgruppe $\mathfrak{G}_1$ n^{ten} Grades kann als lineare homogene Gruppe der gleichen Ordnung in n Variablen dargestellt werden; hingegen kann jede Kollineationsgruppe der Ordnung g in $n-1$ Variablen aus einer homomorphen Gruppe linearer homogener Substitutionen der Determinante $+1$ und der Ordnung ng in n Variablen abgeleitet werden. Man erhält sie, indem man jeder linear gebrochenen Substitution mit der Matrix $\|a_{ik}\|$ $(i, k = 1, 2, \ldots, n)$ die n Matrices $\|\alpha_{ik}^{(j)}\|$ $(i, k, j = 1, 2, \ldots, n)$ entsprechen läßt; hierbei ist $\alpha_{ik}^{(j)} = \varepsilon_j \dfrac{a_{ik}}{\sqrt[n]{|a_{ik}|}}$, wobei ε_j alle n^{ten} Einheitswurzeln durchläuft und $|a_{ik}|$ die Determinante der Matrix $\|a_{ik}\|$ $(i, k = 1, 2, \ldots, n)$ bedeutet.

Zwei Kollineationsgruppen $\mathfrak{G}_1$ und $\mathfrak{G}_2$ heißen *ähnlich, konjugiert, gleichberechtigt* oder *äquivalent*, wenn eine Kollineation P_1 von nicht verschwindender Determinante existiert, so daß $P_1\mathfrak{G}_1P_1^{-1} = \mathfrak{G}_2$ ist.

Die Kollineationsgruppe $\mathfrak{G}_1$ heißt *irreduzibel*, wenn es die zugeordnete homomorphe lineare homogene Substitutionsgruppe ist.

Hat man irgendeine abstrakte endliche Gruppe $\mathfrak{H}$, so kann man sich analog der Aufgabe, alle Darstellungen von $\mathfrak{H}$ durch ganze lineare homogene Substitutionsgruppen zu finden (vgl. S. 229), das Problem stellen, *alle mit $\mathfrak{H}$ homomorphen Kollineationsgruppen* zu bestimmen. Hierzu hat man zuerst eine *Darstellungsgruppe $\mathfrak{D}$ der abstrakten Gruppe $\mathfrak{H}$* aufzusuchen.

Die abstrakte Gruppe $\mathfrak{D}$ heißt eine *Darstellungsgruppe von* $\mathfrak{H}$, wenn sie die Gruppe höchster Ordnung ist, die erstens eine aus invarianten Elementen von $\mathfrak{D}$ bestehende Untergruppe $\mathfrak{M}$ besitzt, so daß die Faktorgruppe $\mathfrak{D}/\mathfrak{M}$ und die Gruppe $\mathfrak{H}$ holoedrisch isomorph sind, zweitens die Kommutatorgruppe von $\mathfrak{D}$ alle Elemente von $\mathfrak{M}$ enthält. Für eine abstrakte Gruppe $\mathfrak{H}$ können auch mehrere nicht isomorphe Darstellungsgruppen existieren. Die endliche kommutative Gruppe $\mathfrak{M}$ ist durch die Gruppe $\mathfrak{H}$ *eindeutig* bestimmt; die Gruppe $\mathfrak{M}$ heißt der *Multiplikator der Gruppe* $\mathfrak{H}$. Hieraus folgt im besonderen, daß alle Darstellungsgruppen einer Gruppe $\mathfrak{H}$ von der gleichen Ordnung sind.

Bestimmt man alle irreduziblen Darstellungen irgend einer abstrakten Darstellungsgruppe $\mathfrak{D}$ *von* $\mathfrak{H}$ *durch ganze lineare homogene Substitutionen und bildet man zu jeder auf diese Weise erhaltenen irreduziblen linearen homogenen Substitutionsgruppe die zugehörige linear gebrochene Gruppe, so ist diese mit* $\mathfrak{H}$ *homomorph. Sieht man ähnliche Kollineationsgruppen als nicht verschieden an, so findet man auf die angegebene Art alle mit* $\mathfrak{H}$ *homomorphen irreduziblen Kollineationsgruppen. Der Grad jeder mit einer abstrakten endlichen Gruppe* $\mathfrak{H}$ *homomorphen irreduziblen Kollineationsgruppe ist ein Divisor der Ordnung von* $\mathfrak{H}$.

Das Problem, alle mit einer abstrakten endlichen Gruppe homomorphen Kollineationsgruppen zu bestimmen, ist von J. Schur, *Journ. f. Math.* **127**, 20 (1904), **132**, 85 (1907) durchgeführt worden. Hierin ist im besonderen die Bestimmung aller mit einer abstrakten Gruppe isomorphen Kollineationsgruppen niedrigsten Grades enthalten (Kleinsches Normalproblem).

Bedenkt man, daß bei jeder Permutation die Summe der Variablen $x_1 + x_2 + \cdots + x_n$ ungeändert bleibt und geht von der sich infolgedessen ergebenden Gruppe linearer homogener Substitutionen in $n-1$ Variablen zu der zugehörigen Kollineationsgruppe in $n-2$ Variablen über, so folgt, daß jede Permutationsgruppe n^{ten} Grades mit einer Kollineationsgruppe $n-1^{\text{ten}}$ Grades, also einer solchen des Raumes R_{n-2} homomorph ist. Für die symmetrischen und alternierenden Permutationsgruppen gilt das Theorem (Wiman, *Math. Ann.* **52**, 243 (1899)):

Die symmetrische Gruppe $\mathfrak{S}_{n!}$ *und die alternierende Gruppe* $\mathfrak{A}_{\frac{n!}{2}}$ *sind mit Kollineationsgruppen* $n-1^{ten}$ *Grades holoedrisch isomorph, nur für die* $\mathfrak{A}_{\frac{4!}{2}}$, *die* $\mathfrak{S}_{4!}$, *die* $\mathfrak{A}_{\frac{5!}{2}}$, *die* $\mathfrak{A}_{\frac{6!}{2}}$, *die* $\mathfrak{S}_{6!}$

und die $\frac{\mathfrak{A}_{71}}{2}$ *existieren holoedrisch isomorphe Kollineationsgruppen niedrigeren Grades. Die* $\frac{\mathfrak{A}_{41}}{2}$ *ist mit der Tetraedergruppe, die* $\mathfrak{S}_{41}$ *mit der Oktaedergruppe und die* $\frac{\mathfrak{A}_{51}}{2}$ *mit der Ikosaedergruppe des* R_1 (vgl. unten), *die* $\frac{\mathfrak{A}_{61}}{2}$ *mit der Valentinergruppe des* R_2 (S. 242), *die* $\mathfrak{S}_{61}$ *und* $\frac{\mathfrak{A}_{71}}{2}$ *mit Kollineationsgruppen des* R_3 *holoedrisch isomorph.* Daß die $\mathfrak{S}_{61}$ und $\frac{\mathfrak{A}_{71}}{2}$ mit Kollineationsgruppen unseres R_3 holoedrisch isomorph sind, vgl. bei Klein (*Math. Ann.* **28**, 499 (1887)). Die Aufzählung der besonderen Kollineationsgruppen des R_2 und R_3, die mit symmetrischen und alternierenden Permutationsgruppen holoedrisch isomorph sind, findet man bei H. Maschke, *Math. Ann.* **51**, 253 (1899).

Betrachtet man ähnliche Kollineationsgruppen als nicht verschieden, so gibt es für $n > 7$ *im* R_{n-2} *nur eine einzige mit* $\mathfrak{S}_{n1}$ *bezw.* $\frac{\mathfrak{A}_{n1}}{2}$ *holoedrisch isomorphe Kollineationsgruppe; eine Ausnahme bildet nur der Fall* $n = 9$, *bei dem im* R_7 *zwei mit der alternierenden Gruppe von 9 Buchstaben isomorphe Kollineationsgruppen existieren.* (Wiman a. a. O.)

Sieht man ähnliche Kollineationsgruppen als nicht verschieden an, so existieren nach dem Jordanschen Satz (vgl. S. 227) *für einen gegebenen Grad n nur eine endliche Anzahl* **endlicher** *Kollineationsgruppen von wesentlich verschiedenem Typus.*

Es gibt fünf verschiedene Typen endlicher binärer **Kollineationsgruppen**, *d. h. solcher des* R_1.

1. Die *zyklische Gruppe* $\mathfrak{C}_g$. Sie hat als Kollineationsgruppe die Ordnung g und geht aus einer linearen homogenen Substitutionsgruppe der Determinante $+1$ hervor, die durch die lineare homogene Substitution S_1 mit der Matrix:

$$\left\| \begin{matrix} e^{\frac{\pi i}{g}} & 0 \\ 0 & e^{-\frac{\pi i}{g}} \end{matrix} \right\|, \qquad e^{2\pi i} = 1$$

erzeugt wird.

2. Die *Diedergruppe* $\mathfrak{D}_{2g}$. Sie hat als Kollineationsgruppe die Ordnung $2g$ und geht aus einer linearen homogenen Sub-

stitutionsgruppe der Determinante $+1$ hervor, die durch zwei lineare homogene Substitutionen

$$S_1 = \begin{Vmatrix} e^{\frac{\pi i}{g}} & 0 \\ 0 & e^{-\frac{\pi i}{g}} \end{Vmatrix}, \quad S_2 = \begin{Vmatrix} 0 & i \\ i & 0 \end{Vmatrix}$$

erzeugt wird.

3. Die *Tetraedergruppe* $\mathfrak{T}_{12}$. Sie hat als Kollineationsgruppe die Ordnung 12 und geht aus einer linearen homogenen Substitutionsgruppe der Determinante $+1$ hervor, die durch drei lineare homogene Substitutionen:

$$S_1 = \begin{Vmatrix} i & 0 \\ 0 & -i \end{Vmatrix}, \quad S_2 = \begin{Vmatrix} 0 & i \\ i & 0 \end{Vmatrix}, \quad S_3 = \begin{Vmatrix} \frac{1-i}{2} & \frac{1-i}{2} \\ -\frac{1+i}{2} & \frac{1+i}{2} \end{Vmatrix}$$

erzeugt wird.

4. Die *Oktaedergruppe* $\mathfrak{O}_{24}$. Sie hat als Kollineationsgruppe die Ordnung 24 und geht aus einer linearen homogenen Substitutionsgruppe der Determinante $+1$ hervor, die durch die zwei linearen homogenen Substitutionen:

$$S_1 = \begin{Vmatrix} \sqrt{i} & 0 \\ 0 & \frac{1}{\sqrt{i}} \end{Vmatrix}, \quad S_2 = \begin{Vmatrix} \frac{1}{\sqrt{2i}} & \frac{1}{\sqrt{2i}} \\ \frac{1}{i\sqrt{2i}} & -\frac{1}{i\sqrt{2i}} \end{Vmatrix}$$

erzeugt wird.

5. Die *Ikosaedergruppe* $\mathfrak{I}_{60}$. Sie hat als Kollineationsgruppe die Ordnung 60 und geht aus einer linearen homogenen Substitutionsgruppe der Determinante $+1$ hervor, die durch die zwei linearen homogenen Substitutionen:

$$S_1 = \begin{Vmatrix} -\varepsilon^3 & 0 \\ 0 & -\varepsilon^2 \end{Vmatrix}, \quad S_2 = \begin{Vmatrix} \frac{1}{\varepsilon - \varepsilon^{-1}} & \frac{1}{\varepsilon^2 - \varepsilon^{-2}} \\ \frac{1}{\varepsilon^2 - \varepsilon^{-2}} & \frac{-1}{\varepsilon - \varepsilon^{-1}} \end{Vmatrix}$$

erzeugt wird; ε ist eine primitive fünfte Einheitswurzel.

Faßt man die obigen fünf Gruppen als Gruppen linearer homogener Substitutionen auf, so enthalten sie sämtlich die Ähnlichkeitssubstitution $\begin{Vmatrix} -1 & 0 \\ 0 & -1 \end{Vmatrix}$ und haben daher doppelt so hohe Ordnungen als die Kollineationsgruppen $\mathfrak{C}_g$, $\mathfrak{D}_{2g}$, $\mathfrak{T}_{12}$, $\mathfrak{O}_{24}$ und $\mathfrak{I}_{60}$. *Nur für die endlichen Gruppen $\mathfrak{C}_g$ bei beliebigem g*

und die $\mathfrak{D}_{2g}$ bei ungeradem g existieren holoedrisch isomorphe lineare homogene Substitutionsgruppen in zwei Variablen.

Für $g = 2$ wird die aus den Kollineationen

$$\xi = \xi', \quad \xi = -\xi', \quad \xi = \frac{1}{\xi'}, \quad \xi = -\frac{1}{\xi'}$$

bestehende Diedergruppe $\mathfrak{D}_4$ als *Vierergruppe* bezeichnet.

Für $g = 3$ ist die Diedergruppe $\mathfrak{D}_6$ ähnlich mit der *anharmonischen Gruppe*. Diese liefert in der projektiven Geometrie die 6 zusammengehörigen Werte

$$\xi = \xi', \ \xi = \frac{1}{\xi'}, \ \xi = 1 - \xi', \ \xi = \frac{1}{1-\xi'}, \ \xi = \frac{\xi'-1}{\xi'}, \ \xi = \frac{\xi}{\xi'-1}$$

für das anharmonische Doppelverhältnis von vier Punkten einer Geraden und ferner die 6 zusammengehörigen Werte des Legendreschen Moduls k^2 eines elliptischen Integrals erster Gattung.

$\mathfrak{T}_{12}$ und $\mathfrak{J}_{60}$ sind mit den alternierenden Permutationsgruppen in 4 bezw. 5 Symbolen holoedrisch isomorph. Die $\mathfrak{O}_{24}$ ist mit der symmetrischen Permutationsgruppe in 4 Symbolen holoedrisch isomorph.

Die Kollineationsgruppen $\mathfrak{D}_{2g}$, $\mathfrak{T}_{12}$, $\mathfrak{O}_{24}$ und $\mathfrak{J}_{60}$ sind mit den endlichen Gruppen holoedrisch isomorph, welche die regulären Körper „Dieder, Tetraeder, Oktaeder und Ikosaeder" mit sich zur Deckung bringen (vgl. S. 176); hierbei gilt das reguläre g-Eck, *Dieder* genannt, auch als regulärer Körper. Die endlichen Kollineationsgruppen des R_1 werden daher auch als *Polyedergruppen* bezeichnet.

Sämtliche Substitutionen A der linearen homogenen homomorphen endlichen Substitutionsgruppen der Determinante $+1$ des R_1, aus denen die endlichen Kollineationsgruppen hervorgehen, lassen sich in die Form bringen:

$$x_1 = (a + ib)x_1' + (ci + d)x_2',$$
$$x_2 = (ci - d)x_1' + (a - ib)x_2',$$

wobei $a^2 + b^2 + c^2 + d^2 = 1$, $i = \sqrt{-1}$ und a, b, c, d reelle Größen bedeuten; dies folgt aus der symbolischen Gleichung $\bar{A}' = A^{-1}$ (vgl. S. 134 letzte Zeile), die besagt, daß alle endlichen Gruppen linearer homogener Substitutionen in zwei Variablen so transformierbar sind, daß sie die Hermitesche Form $x_1\bar{x}_1 + x_2\bar{x}_2$ in sich überführen. Die entsprechenden Kollineationen $\xi = \dfrac{(a+ib)\xi' + (ci+d)}{(ci-d)\xi' + (a-ib)}$ stellen die sämtlichen

Drehungen einer Kugel um ihren Durchmesser dar, wenn man die Kugel nach Riemann als Trägerin der komplexen Variablen ξ interpretiert (Cayley, *Math. Ann.* **16**, 260 (1880), Klein, *Ikosaeder*, S. 34).

Ein *volles Invariantensystem* für die oben angegebene lineare homogene Substitutionsgruppe, aus der die Diedergruppe $\mathfrak{D}_{2g}$ hervorgeht, bilden die drei Formen:

$$F_1 = x_1 x_2,\ F_2 = x_1^g + x_2^g,\ F_3 = x_1^g - x_2^g.$$

Sie sind sämtlich relative Invarianten der Gruppe. Zwischen ihnen besteht die Relation $4F_1^{2g} = F_2^2 - F_3^2$.

Ein volles Invariantensystem für die lineare homogene Substitutionsgruppe, aus der die Tetraedergruppe $\mathfrak{T}_{12}$ hervorgeht, bilden die Formen:

$$f_4 = x_1^4 + 2\sqrt{-3}\, x_1^2 x_2^2 + x_2^4,$$
$$f_4' = x_1^4 - 2\sqrt{-3}\, x_1^2 x_2^2 + x_2^4,$$
$$f_6 = x_1 x_2 (x_1^4 - x_2^4).$$

Zwischen ihnen besteht die Relation:

$$12\sqrt{-3}\, f_6^2 = f_4^3 - f_4'^3.$$

f_6 ist eine absolute Invariante der Gruppe, f_4 und f_4' sind relative Invarianten vom Index 3. Ein volles System absoluter Invarianten bilden f_6, $H_8 = f_4 \cdot f_4'$, f_4^3.

Ein volles Invariantensystem für die lineare homogene Substitutionsgruppe, aus der die Oktaedergruppe $\mathfrak{O}_{24}$ hervorgeht, bilden die drei Formen:

$$f_6 = x_1 x_2 (x_1^4 - x_2^4),$$
$$H_8 = f_4 f_4' = x_1^8 + 14 x_1^4 x_2^4 + x_2^8,$$
$$K = \tfrac{1}{2}(f_4^3 + f_4'^3) = x_1^{12} - 33 x_1^8 x_2^4 - 33 x_1^4 x_2^8 + x_2^{12}.$$

Zwischen ihnen besteht die Relation:

$$108 f_6^4 = H_8^3 - K^2.$$

Die Form H_8 ist eine absolute Invariante, f_6 und K sind relative Invarianten vom Index 2. Ein volles System absoluter Invarianten bilden die drei Formen H_8, f_6^2 und $f_6 \cdot K$.

Ein volles Invariantensystem für die lineare homogene Substitutionsgruppe, aus der die Ikosaedergruppe hervorgeht, bilden die drei Formen:

$$f_{12} = x_1 x_2 (x_1^{10} + 11 x_1^5 x_2^5 - x_2^{10}),$$
$$H_{20} = -(x_1^{20} + x_2^{20}) + 228(x_1^{15} x_2^5 - x_2^{15} x_1^5) - 494 x_1^{10} x_2^{10},$$
$$T = x_1^{30} + x_2^{30} + 522(x_1^{25} x_2^5 - x_1^5 x_2^{25}) - 10005(x_1^{20} x_2^{10} + x_2^{20} x_1^{10}).$$

Zwischen ihnen besteht die Relation:

$$T^2 + H_{20}^3 = 1728 f_{12}^5.$$

Die drei Formen f_{12}, H_{20} und T sind absolute Invarianten; für die Ikosaedergruppe existieren keine relativen Invarianten.

Die Formen f_4, f_4', f_6 und f_{12} und die aus ihnen durch lineare homogene Transformation hervorgehenden sind die einzigen binären Formen ohne vielfache Faktoren, deren vierte Überschiebung $(f,f)^4$ verschwindet. (Wedekind, *Habilit.-Schr.*, Karlsruhe (1876), Brioschi, *Ann. di mat.* (2) 8, 24 (1877), Gordan, *Vorlesungen über Invariantentheorie* 2, 204.)

H_8 ist die Hessesche Determinante von f_6, H_{20} die von f_{12}, f_4' die von f_4 und f_4 die von f_4', f_6 ist die Funktionaldeterminante von f_4 und f_4', K von f_6 und H_8, T von H_{20} und f_{12}.

Die Form f_4 entspricht den Ecken des Tetraeders, f_4' denjenigen des Gegentetraeders, f_6 den Kantenhalbierungspunkten, die ein Oktaeder bilden. Ebenso liegt es bei den Invarianten des Oktaeders und des Ikosaeders; f_6 entspricht den Oktaederecken, H_8 dem Polarwürfel, K den 12 Kantenhalbierungspunkten, f_{12} entspricht den Ikosaederecken, H_{20} den Ecken des reziproken Pentagondodekaeders und T den Kantenhalbierungspunkten. Die Hesseschen Determinanten entsprechen also den Mitten der Seitenflächen der betrachteten Polyeder, die Funktionaldeterminanten den Mittelpunkten der Kanten der Polyeder. Setzt man $\frac{x_1}{x_2} = \vartheta$ und bildet sich $\frac{f_4^3}{f_4'^3} = Z$, $\frac{H_8^3}{108 f_6^4} = Z$, $\frac{H_{20}^3}{1728 f_{12}^5} = Z$, so ist Z Funktion von ϑ. Die drei auf diese Weise definierten Gleichungen $f(\vartheta) = Z$ vom Grade 12, 24 und 60 heißen nach Klein die *Tetraeder-*, *Oktaeder-* und *Ikosaedergleichung*. Die letztere Gleichung lautet ausführlich:

$$(-\vartheta^{20} - 1 + 228\vartheta^{15} - 228\vartheta^5 - 494\vartheta^{10})^3 - 1728\vartheta^5 Z(\vartheta^{10} + 11\vartheta^5 - 1)^5 = 0.$$

ϑ als Funktion von Z aufgefaßt, heißt die Irrationalität des Tetraeders bezw. des Oktaeders und Ikosaeders.

Wir vermerken noch die *Untergruppen der Ikosaedergruppe.* Sie sind zyklische Gruppen: fünfzehn $\mathfrak{C}_2$, zehn $\mathfrak{C}_3$, sechs $\mathfrak{C}_5$, Diedergruppen: fünf $\mathfrak{D}_4$, zehn $\mathfrak{D}_6$, sechs $\mathfrak{D}_{10}$ und 5 Tetraedergruppen $\mathfrak{T}_{12}$. Gruppen gleicher Ordnung sind in bezug auf die Ikosaedergruppe konjugiert.

Die endlichen Kollineationsgruppen des R_1 sind zuerst von F. Klein, *Math. Ann.* **9**, 183 (1876), und zwar auf geometrischem Wege durch Betrachtung der Drehungen einer Kugel um ihren Durchmesser und der der Kugel einbeschriebenen Polyeder bestimmt worden. Die erste algebraische Herleitung der endlichen Kollineationsgruppen des R_1 gab Gordan, *Math. Ann.* **12**, 23 (1877). Auf die invarianten binären Formen in unhomogener Gestalt war bereits H. A. Schwarz, *Journ. f. Math.* **75**, 323 (1873) bei der Bestimmung der Fälle gelangt, in denen die Gaußsche Differentialgleichung der hypergeometrischen Funktion durch algebraische Funktionen befriedigt wird. Vgl. die systematischen Darstellungen bei Klein, *Ikosaeder*, H. Weber, *Algebra* **2**, 269, Bianchi, *Teoria dei gruppi,* S. 99, Vivanti, *Elementi della teoria delle funzioni poliedriche e modulari,* Milano 1906.

Die Aufzählung der endlichen ternären Kollineationsgruppen, d. h. derjenigen des R_2 hat zuerst C. Jordan (*Journ. f. Math.* **84**, 89 (1878)) unternommen; hierbei entgingen ihm eine von F. Klein (*Math. Ann.* **14**, 428 (1879)) entdeckte Gruppe der Ordnung 168 und eine von Valentiner (*Videnskabernes Selskabs Skrifter,* 6. Raekke (1889), Kopenhagen) gefundene Gruppe der Ordnung 360. Eine erneute Bestimmung der endlichen ternären Kollineationsgruppen gab Blichfeldt *Math. Ann.* **63**, 552 (1907).

Abgesehen von reduziblen Kollineationsgruppen und solchen, bei denen sich die Variablen der zugehörigen homogenen linearen Substitutionsgruppen bis auf konstante Faktoren nur untereinander permutieren (*Monomialgruppen,* Maschke, *Am. J. math.* **17**, 168 (1895)), *gibt es, wenn ähnliche Kollineationsgruppen als nicht verschieden gelten, sechs verschiedene Arten endlicher ternärer Kollineationsgruppen, d. h. solcher der Ebene.*

1. Die *Hessesche Gruppe*, die als Kollineationsgruppe die *Ordnung 216* hat und aus keiner niedrigeren ternären linearen homogenen Substitutionsgruppe als einer solchen der Ordnung 648 herleitbar ist. Der Name der Gruppe stammt daher, weil sie mit der Hesseschen Konfiguration zusammenhängt, die sich bei der Betrachtung des syzygetischen Büschels $kf + \lambda\Delta = 0$ ergibt, der von einer Kurve dritter Ordnung $f = 0$ und ihrer

Hesseschen Kurve Δ gebildet wird. Eine Invariante der Gruppe ist die Form, welche die vier Wendepunktdreiecke des Büschels bestimmt. Das volle Invariantensystem, das aus fünf Formen besteht, findet man bei Maschke, *Math. Ann.* **33**, 326 (1889).

2. *Eine invariante Untergruppe der Hesseschen Gruppe von der Ordnung 72* bezw. 216 bei homogener Schreibweise.

3. *Eine Untergruppe der Hesseschen Gruppe von der Ordnung 36* bezw. 108 bei homogener Schreibweise. Die Hessesche Gruppe enthält drei solcher Gruppen; sie haben eine in der Hesseschen Gruppe der Ordnung 216 invariante Untergruppe der Ordnung 18 gemein. Die Hessesche Gruppe ist daher mit der Tetraedergruppe $\mathfrak{T}_{12}$ homomorph; nach ihr transformieren sich die Formen f und Δ.

4. Die *ternäre Ikosaedergruppe,* die als Kollineationsgruppe die Ordnung 60 hat und aus einer linearen homogenen ternären Substitutionsgruppe derselben Ordnung 60 hervorgeht.

5. Die *Kleinsche Gruppe,* die als Kollineationsgruppe die Ordnung 168 hat und aus einer linearen homogenen ternären Substitutionsgruppe derselben Ordnung 168 hervorgeht. Die Gruppe ist *einfach* und, da es nur eine einfache abstrakte Gruppe der Ordnung 168 gibt, holoedrisch isomorph mit der Modulargruppe für die Primzahl $p = 7$ (vgl. S. 201). Die Invariante niedrigsten Grades der homogenen Gruppe ist die ternäre Form

$$x_1^3 x_3 + x_2^3 x_1 + x_3^3 x_2 .$$

Die Gruppe besitzt nur absolute Invarianten. Das volle Invariantensystem besteht aus vier Formen. Die Gruppe ist gefunden von Klein, *Math. Ann.* **14**, 428 (1879), vgl. ferner Klein, *Math. Ann.* **15**, 265 (1879), Gordan, ebenda **17**, 217, 359 (1880), **20**, 515 (1882), **25**, 459 (1885). Systematische Darstellungen bei Klein-Fricke, *Modulfunktionen* **1**, 692 u. H. Weber, *Algebra* **2**, 497.

6. Die *Valentinersche Gruppe,* die als Kollineationsgruppe die Ordnung 360 hat und aus keiner niedrigeren homogenen linearen ternären Gruppe als einer solchen der Ordnung 1080 hervorgeht. *Die Valentinersche Gruppe ist mit der alternierenden Permutationsgruppe in sechs Symbolen holoedrisch isomorph.* Sie enthält zwei Systeme von je sechs konjugierten Ikosaedergruppen und zwei Systeme von je 15 konjugierten Oktaedergruppen. Die Invariante niedrigsten Grades der homogenen Gruppe ist von der Form:

$$10 x_1^3 x_2^3 + 9 x_3 (x_1^5 + x_2^5) - 45 x_1^2 x_2^2 x_3^2 - 135 x_1 x_2 x_3^4 + 27 x_3^6 .$$

Das volle Invariantensystem der Gruppe besteht aus vier Formen und ist zuerst von Wiman, *Math. Ann.* **47**, 531 (1896) aufgestellt worden. Vgl. ferner Gerbaldi, *Rend. Circolo di Palermo,* **12**, 23 (1898), **13**, 161 (1899), **14**, 66 (1900), **16**, 129 (1902), *Math. Ann.* **50**, 473 (1898), Fricke, *Gött. Nachr.* (1896), 199, *Math. Ver.* **5**, 55 (1896), L. Lachtin, *Math. Ann.* **51**, 463 (1899), Gordan, *Math. Ann.* **61**, 453 (1905).

Die vollständige Aufzählung der verschiedenen endlichen quaternären Kollineationsgruppen, d. h. derjenigen unseres R_3, stammt von Blichfeldt, *Math. Ann.* **60**, 204 (1905), *Trans. Am. M. S.* **6**, 230 (1905). Vgl. ferner Autonne, *Journ. de math.* (5) **7**, 351 (1901) u. Bagnera, *Rend. Circolo di Palermo* **19**, 1 (1905). Im R_3 gibt es allein 25 nicht ähnliche Kollineationsgruppen, die mit symmetrischen und alternierenden Permutationsgruppen holoedrisch isomorph sind (Maschke, *Math. Ann.* **51**, 298 (1899)), hierunter je eine mit $\mathfrak{S}_{61}$ und $\mathfrak{A}_{71}$ isomorphe Gruppe. Besonders bemerkenswert sind unter den quaternären endlichen Kollineationsgruppen:

1. eine solche der Ordnung 11520, die aus einer homogenen quaternären linearen Substitutionsgruppe von viermal so hoher Ordnung hervorgeht,

2. eine solche der Ordnung 25920, die sich aus einer homogenen quaternären linearen Substitutionsgruppe von doppelt so hoher Ordnung ergibt.

Nach der Kollineationsgruppe der Ordnung 11520, die F. Klein (*Math. Ann.* **2**, 198 (1870), **4**, 356 (1871)) durch liniengeometrische Betrachtungen gefunden hat, transformieren sich die Verhältnisse der bei den hyperelliptischen Funktionen des Geschlechtes $p = 2$ auftretenden sogenannten Borchardtschen Moduln. Die $\mathfrak{G}_{11520}$ ist holoedrisch isomorph mit der Galoisschen Gruppe der Gleichung 16ten Grades, welche die 16 Knotenpunkte einer Kummerschen Fläche bestimmt (Jordan, *Traité,* 313) und homomorph mit der symmetrischen Gruppe $\mathfrak{S}_{61}$. Ein volles Invariantensystem der homogenen Gruppe ist von Maschke, *Math. Ann.* **30**, 496 (1887) aufgestellt worden. Weitere Literatur: Reichardt, *Math. Ann.* **28**, 84 (1887), *Nova Acia Leop.* **50**, 375 (1887), Heß, ebenda **55**, 100 (1891).

Die Kollineationsgruppe $\mathfrak{G}_{25920}$ ist *einfach.* Sie ist holoedrisch isomorph mit einer invarianten Untergruppe des Index 2 der Galoisschen Gruppe $\mathfrak{G}_{51840}$ der Ordnung 51840 der Gleichung für die Bestimmung der 27 Geraden einer allgemeinen Fläche

dritter Ordnung; von dieser $\mathfrak{G}'_{51840}$ hängt auch die Dreiteilung der hyperelliptischen Funktionen ab. (C. Jordan, *Traité*, 316 u. 365, F. Klein, *Journ. de math.* (4) **4**, 169 (1888), H. Burkhardt, *Math. Ann.* **38**, 161 (1891), **41**, 313 (1893).) Das volle Invariantensystem der homogenen $\mathfrak{G}_{51840}$[1]), die mit der Kollineationsgruppe $\mathfrak{G}_{25940}$ homomorph ist, bei H. Maschke, *Math. Ann.* **33**, 317 (1889). Eine gruppentheoretische Untersuchung dieser vielfach behandelten Gruppe mit Literaturangaben bei Dickson, *Trans. Am. M. S.* **5**, 126 (1904), *Proc. Lond. M. S.* (2) **1**, 283 (1904), vgl. auch Dickson, *Linear groups*, S. 303, W. Burnside, *Proc. Royal Soc.* A **77**, 182 (1906). Vgl. auch S. 249.

Erwähnt sei noch, daß nach den Untersuchungen von J. Schur, *Journ. f. Math.* **132**, 135 (1907) über die Darstellungen der verallgemeinerten Modulargruppe als Kollineationsgruppen die einfache Gruppe der Ordnung 504 sich als Kollineationsgruppe in sechs und nicht weniger Variablen darstellen läßt, während die voraufgehenden einfachen Gruppen der Ordnungen 60 und 168 als Kollineationsgruppen in einer bezw. zwei Variablen darstellbar sind. Es gilt folgender allgemeiner Satz: Ist $p^m > 3$, aber von 9 verschieden, so gibt es, wenn p eine ungerade Primzahl ist, im ganzen $\frac{p^m+3}{2}$ verschiedene irreduzible Gruppen ganzer linearer homogener Substitutionen, die *mit der verallgemeinerten endlichen Modulargruppe* (vgl. S. 215) *der Ordnung* $\frac{p^m(p^{2m}-1)}{2}$ *isomorph sind,* und zwar eine des Grades p^m, $\frac{p^m-4-\varepsilon}{4}$ des Grades p^m+1, $\frac{p^m-2+\varepsilon}{4}$ des Grades p^m-1 und zwei des Grades $\frac{p^m+\varepsilon}{2}$. Ebenso gibt es für $p^m > 3$, aber $\neq 9$, $\frac{p^m+3}{2}$ verschiedene irreduzible Kollineationsgruppen, die mit der verallgemeinerten Modulargruppe isomorph sind und sich nicht als Gruppen von $\frac{p^m(p^{2m}-1)}{2}$ ganzen linearen homogenen Substitutionen schreiben lassen. Unter diesen Gruppen haben $\frac{p^m-2+\varepsilon}{4}$ den Grad p^m+1, $\frac{p^m-\varepsilon}{4}$ den Grad p^m-1 und zwei den Grad $\frac{p^m-\varepsilon}{2}$; hierbei ist $\varepsilon = (-1)^{\frac{p^m-1}{2}}$. Es gibt 2^m verschiedene irreduzible Kollineationsgruppen, die mit

1) $\mathfrak{G}_{51840}$ und $\mathfrak{G}'_{51840}$ sind natürlich nicht isomorph; die erste Gruppe hat eine invariante Untergruppe zweiter Ordnung, die zweite eine invariante Untergruppe $\mathfrak{G}_{25920}$.

der *verallgemeinerten endlichen Modulargruppe der Ordnung* $2^m(2^{2m}-1)$ *des* $GF[2^m]$ *für* $2^m > 4$ *isomorph* sind, und zwar hat man eine Gruppe des Grades 2^m, ferner $2^{m-1}-1$ des Grades 2^m+1 und 2^{m-1} des Grades 2^m-1. Alle diese Gruppen lassen sich auch als ganze lineare homogene Substitutionsgruppen derselben Ordnung $2^m(2^{2m}-1)$ darstellen. (J. Schur, a. a. O., 133 u. 134.) Die mit der gewöhnlichen Modulargruppe $(m=1)$ isomorphen Kollineationsgruppen des Grades $\frac{p-1}{2}$ bezw. $\frac{p+1}{2}$, d. h. des $R_{\frac{p-3}{2}}$ bezw. $R_{\frac{p-1}{2}}$, hat F. Klein, *Math. Ann.* **15**, 275 (1879) aus der Transformationstheorie der elliptischen Funktionen gefunden.

§ 12. Lineare Gruppen mit Koeffizienten aus einem Körper. Kongruenzgruppen.

Die Gesamtheit aller Kollineationen A_1:

$$\xi_i = \frac{a_{i1}\xi_1' + a_{i2}\xi_2' + \cdots + a_{in}\xi_n' + a_{in+1}}{a_{n1}\xi_1' + a_{n2}\xi_2' + \cdots + a_{nn}\xi_n' + a_{nn+1}} \quad (i=1,2,\ldots,n)$$

gleichen Grades $n+1$ mit nicht verschwindenden Determinanten $|a_{ik}|$ $(i,k=1,2,\ldots,n+1)$, deren Koeffizienten sämtlich ausnahmslos einem Körper Ω angehören, bildet eine Gruppe; sie heißt die *allgemeine projektive oder linear gebrochene Gruppe in n Variablen* oder die *allgemeinste Kollineationsgruppe* $(n+1)^{ten}$ *Grades mit Koeffizienten aus* Ω. Ist Ω ein endlicher Körper (vgl. S. 177), so ist die Gruppe selbst *endlich*; man spricht von der *allgemeinen projektiven oder linear gebrochenen* oder *allgemeinsten Kollineationsgruppe mit Koeffizienten aus dem* $GF[p^m]$ (Galoissches Feld der Ordnung p^m) oder kürzer von einer *allgemeinen linear gebrochenen Kongruenzgruppe.*

Die Gesamtheit aller Substitutionen

$$\xi_i = a_{i1}\xi_1' + a_{i2}\xi_2' + \cdots + a_{in}\xi_n' + a_{in+1} \quad (i=1,2,\ldots,n)$$

gleichen Grades mit nicht verschwindenden Determinanten $|a_{ik}|$ $(i,k=1,2,\ldots,n)$, deren Koeffizienten ausnahmslos einem Körper Ω angehören, bildet die *allgemeine lineare unhomogene Gruppe* oder die *allgemeine volle lineare Gruppe mit Koeffizienten aus* Ω. Sie hat die *kommutative Gruppe*: $\xi_i = \xi_i' + a_{in+1}$ $(i=1,2,\ldots,n)$ zur *invarianten Untergruppe*. Die letztere Gruppe wird auch, wenn Ω das $GF[p^m]$ ist, als *arithmetische Gruppe* bezeichnet.

In der allgemeinen linearen unhomogenen Gruppe mit Koeffizienten aus Ω ist die *allgemeine lineare homogene Gruppe*

mit Koeffizienten aus Ω als Untergruppe enthalten. Sie besteht aus der Gesamtheit aller Substitutionen:

$$\xi_i = a_{i1}\xi_1' + a_{i2}\xi_2' + \cdots + a_{in}\xi_n' \qquad (i = 1, 2 \ldots, n)$$

mit nicht verschwindenden Determinanten $|a_{ik}|$ $(i, k = 1, 2, \ldots, n)$ und Koeffizienten aus Ω.

Ist Ω das $GF[p^m]$, so hat die allgemeine lineare homogene Kongruenzgruppe n^{ten} Grades[1]) *die Ordnung*

$$h = (p^{mn} - 1)(p^{mn} - p^m) \cdot (p^{mn} - p^{2m}) \cdots (p^{mn} - p^{m(n-1)})$$

und die volle Gruppe die Ordnung $h \cdot p^{mn}$

Läßt man in dem Symbol $l_{\xi_1'\xi_2'\ldots\xi_n'}$ die n Indices die p^m Marken des $GF[p^m]$ durchlaufen und ersetzt ξ_i' durch

$$\xi_i = a_{i1}\xi_1' + a_{i2}\xi_1' + \cdots + a_{in}\xi_n' + a_{in+1} \qquad (i = 1, 2, \ldots, n),$$

wobei die Koeffizienten a_{ik} alle Zahlen des $GF[p^m]$ mit nicht verschwindenden Determinanten $|a_{ik}|$ $(i, k = 1, 2, \ldots, n)$ durchlaufen, so entspricht jeder linearen unhomogenen Substitution eine Permutation unter den p^{mn} Größen $l_{\xi_1'\xi_2'\ldots\xi_n'}$. Die aus diesen Permutationen bestehende Gruppe ist mit der linearen unhomogenen Substitutionsgruppe mit Koeffizienten aus dem $GF[p^m]$ holoedrisch isomorph. Anstatt ein Symbol l mit n Indices zu verwenden, kann man auch ein Symbol l_ζ mit nur einem Index ζ verwenden, den man die p^{nm} Marken des $GF[p^{mn}]$ durchlaufen läßt. Ersetzt man ζ durch eine geeignet gewählte ganze rationale Funktion $\psi(\zeta)$ von ζ mit Koeffizienten aus dem $GF[p^{mn}]$, so stellt $\begin{pmatrix} l_\zeta \\ l_{\psi(\zeta)} \end{pmatrix}$, wenn ζ die Zahlen des $GF[p^{mn}]$ durchläuft, eine Permutation unter den p^{mn} Größen l_ζ dar. Auch eine aus derartigen Permutationen bestehende Permutationsgruppe des $p^{mn\,ten}$ Grades läßt sich der linearen unhomogenen Substitutionsgruppe mit Koeffizienten aus dem $GF[p^m]$ holoedrisch isomorph zuordnen. (Betti, *Ann. di scienze mat. e fis.* 3 (1852), 6 (1855), *Opere mat.*, Milano 1903, 1, 48 u. 117, Mathieu, *Journ. de math.* (2) 6, 301 (1861), Dickson, *Linear groups*, S. 64 u. 75.)

Die sich für $m = 1$ ergebende Permutationsgruppe tritt bereits in folgendem berühmtem Satze von Galois (*Œuvres*, p. 27) auf:

---

1) Unter dem Grad einer linearen homogenen Gruppe versteht man die Variablenzahl. Eine Kongruenzgruppe erhält man auch, wenn man alle linearen homogenen Substitutionen gleichen Grades mit ganzzahligen Koeffizienten mod s betrachtet, deren Determinanten zu s relativ prim sind, wobei s eine beliebig gewählte, ganze positive Zahl bedeutet. (C. Jordan, *Traité*, S. 92.)

Jede primitive Permutationsgruppe des Grades p^n (p Primzahl) kann nur dann auflösbar sein, wenn sie eine Untergruppe der Permutationsgruppe $p^{n\,ten}$ Grades der Ordnung

$$(p^n - 1)(p^n - p)(p^n - p^2) \dots (p^n - p^{n-1})p^n$$

ist, die mit der allgemeinen vollen linearen Kongruenzgruppe mit Koeffizienten mod p in n Variablen holoedrisch isomorph ist. Für $n = 1$ ist die Gruppe immer auflösbar, vgl. S. 213. Für $n > 1$ ist die Gruppe der Ordnung $(p^n - 1)(p^n - p)(p^n - p^2) \dots (p^n - p^{n-1})p^n$ nicht stets auflösbar; mit der Bestimmung ihrer sämtlichen auflösbaren Untergruppen beschäftigt sich C. Jordan, *Traité*, S. 398ff. Von Lehrbüchern vgl. H. Weber, *Algebra* 2, 359. Für $p = 3$, $n = 2$ hat die Gruppe die Ordnung 432. Diese Gruppe ist eine auflösbare Gruppe und die Galoissche Gruppe der *Gleichung neunten Grades,* von der die Bestimmung der *9 Wendepunkte einer Kurve dritter Ordnung* abhängt (Hesse, *Journ. f. Math.* 34 (1847), *Ges. Werke*, 137; vgl. auch H. Weber, *Algebra* 2, 410, Netto, *Algebra* 2, 460).

Die Gesamtheit aller linear gebrochenen Substitutionen:

$$\xi_i = \frac{a_{i1}\xi_1' + a_{i2}\xi_2' + \dots + a_{in}\xi_n' + a_{in+1}}{a_{n1}\xi_1' + a_{n2}\xi_2' + \dots + a_{nn}\xi_n' + a_{nn+1}} \qquad (i = 1, 2, \dots, n)$$

mit Koeffizienten aus einem Körper Ω, bei denen die Determinante $|a_{ik}|$ $(i, k = 1, 2 \dots, n+1)$ den Wert $+1$ hat, heißt die *spezielle Kollineationsgruppe $n+1^{ten}$ Grades* oder die *spezielle linear gebrochene Gruppe oder die spezielle projektive Gruppe mit Koeffizienten aus Ω.* Sie ist eine invariante Untergruppe der allgemeinen projektiven Gruppe mit Koeffizienten aus Ω.

Ist Ω das $GF[p^m]$, so ist die spezielle linear gebrochene Gruppe in n Variablen endlich und hat die Ordnung

$$\frac{1}{d}(p^{m(n+1)} - 1)p^{mn}(p^{mn} - 1)p^{m(n-1)} \dots (p^{2m} - 1)p^m,$$

wobei d der größte gemeinsame Teiler von $n+1$ und $p^m - 1$ ist. Ausgenommen in den zwei Fällen $(n+1, p^m) = (2, 2)$ und $(2, 3)$ ist die endliche spezielle linear gebrochene Gruppe eine einfache Gruppe. Ist die Zahl n im besonderen gleich 1, so erhält man die verallgemeinerte Modulargruppe der Ordnung $(2^{2m} - 1) \cdot 2^m$ bezw. $\frac{(p^{2m} - 1)}{2} p^m$, falls $p > 2$. Für $n = 2$, $p = 2$, $m = 1$ hat die spezielle lineare gebrochene Gruppe in 2 Variablen die Ordnung 168. Als einfache Gruppe muß sie, da es nur eine einfache Gruppe der Ordnung 168 gibt, mit der Modulargruppe

für $p = 7$ holoedrisch isomorph sein. Über die spezielle linear gebrochene Gruppe für $n = 2$, p^m bei beliebigem m vgl. Dickson, *Linear groups*, S. 242.

Die Gesamtheit aller linearen homogenen Substitutionen

$$\xi_i = a_{i1}\xi_1' + a_{i2}\xi_2' + \cdots + a_{in}\xi_n' \qquad (i = 1, 2 \ldots, n)$$

mit Koeffizienten aus einem Körper Ω, bei denen die Determinante $|a_{ik}|$ den Wert $+1$ hat, heißt die *spezielle lineare homogene Gruppe mit Koeffizienten aus* Ω. Die Gruppe hat die aus den Ähnlichkeitstransformationen $\xi_i = \nu\xi_i'$ $(\nu^n = 1)$ gebildete Gruppe zur invarianten Untergruppe, wobei ν alle Größen aus Ω durchläuft.

Ist Ω ein *algebraischer Zahlkörper*, so bildet die Gesamtheit aller linearen homogenen Substitutionen der Determinante $+1$, deren Koeffizienten ausnahmslos *ganze algebraische Zahlen* des Körpers sind, eine Gruppe. Diese Gruppe wird von einer *endlichen Anzahl von Elementen erzeugt.* (Vgl. S. 193.) (A. Hurwitz, *Gött. Nachr.* (1895), 332.) Im besonderen wird also auch die Gruppe aller ganzzahligen linearen homogenen Substitutionen der Determinante $+1$ von einer endlichen Anzahl von Substitutionen erzeugt; für diesen speziellen Fall genügen sogar zwei erzeugende Substitutionen (Kronecker, *Monatsb. d. Berl. Akad.* (1866), *Ges. Werke* **1**, 159, Krazer, *Ann. di mat.* (2) **12**, 283 (1884), W. Burnside, *Messenger of mat.* **33**, 133 (1904)).

Als Untergruppe der speziellen linearen homogenen Gruppe sei für $n = 2r$ die *spezielle lineare Abelsche Gruppe* angeführt. (Die Bezeichnung hat mit kommutativer Gruppe (vgl. S. 174) nichts zu tun; die Gruppe stammt aus der Theorie der Abelschen Funktionen (Jordan, *Traité*, S. 171)). Die spezielle lineare Abelsche Gruppe besteht aus der Gesamtheit aller linearen homogenen Substitutionen:

$$\xi_i = a_{i1}\xi_1' + a_{i2}\xi_2' + \cdots + a_{in}\xi_n', \qquad (i = 1, 2, \ldots, 2r)$$

welche die alternierende bilineare Form:

$$\xi_1\eta_2 - \xi_2\eta_1 + \xi_3\eta_4 - \xi_4\eta_2 + \cdots + \xi_{2r-1}\eta_{2r} - \xi_{2r}\eta_{2r-1}$$

mit kogredienten Variablen ξ_i, η_i in sich überführen. *Ist* Ω *das* $GF[p^m]$, *so hat die sich aus der speziellen linearen Abelschen Gruppe ergebende Gruppe linear gebrochener Substitutionen die Ordnung*

$$\frac{1}{a}(p^{m(2r)} - 1) \cdot p^{m(2r-1)} \cdot (p^{m(2r-2)} - 1)p^{m(2r-3)} \ldots (p^{2m} - 1)p^m,$$

wobei $a = 1$ oder $= 2$, je nachdem $p = 2$ oder > 2 ist. Außer für die drei Fälle $p^m = 2$, $r = 1$ und $r = 2$, $p^m = 3$, $r = 1$ ist die spezielle linear gebrochene Abelsche Gruppe einfach (vgl. S. 202 erste Zeile). Die einfache Gruppe niedrigster Ordnung, die in diesem System enthalten ist, ist eine solche der Ordnung 25920, die $p^m = 3$, $r = 2$ entspricht und mit der einfachen Kollineationsgruppe des R_3 (vgl. S. 243) holoedrisch isomorph ist. $p^m = 2$, $r = 3$ ergibt die einfache Gruppe der Ordnung 1451520; sie ist holoedrisch isomorph mit der Galoisschen Gruppe der *Gleichung der 28 Doppeltangenten einer Kurve vierter Ordnung ohne Doppelpunkte.* Eine Untergruppe der Gruppe $\mathfrak{G}_{1451520}$ vom Index 28 ist holoedrisch isomorph mit der Galoisschen Gruppe $\mathfrak{G}'_{51840}$ (vgl. S. 244) der Gleichung für die 27 Geraden einer Fläche dritter Ordnung. (Jordan, *Traité,* S. 229 u. 330, H. Weber, *Algebra* **2**, 419.) Vgl. auch Dickson, *Trans. Am. M. S.* **3**, 38 u. 377 (1902), *Am. J. math.* **26**, 243 (1904).

Die Gesamtheit aller linearen homogenen Substitutionen

$$\xi_i = a_{i1}\xi_1' + a_{i2}\xi_2' + \cdots + a_{in}\xi_n' \qquad (i = 1, 2, \ldots, n),$$

mit Koeffizienten aus einem Körper Ω, die $\xi_1^2 + \xi_2^2 + \cdots + \xi_n^2$ in sich transformieren, bildet die *orthogonale Gruppe mit Koeffizienten aus Ω.* Ihre Behandlung für das $GF[p^m]$ bei Dickson, *Linear groups,* S. 156. Wegen weiterer Einzelheiten in bezug auf Kongruenzgruppen vgl. das zitierte Werk, sowie in Ergänzung für beliebige Körper: Dickson, *Trans. Am. M. S.* **2**, 363 (1901).

In Analogie mit dem in § 10 behandelten Problem der Darstellung einer abstrakten Gruppe als Gruppe linearer homogener Substitutionen kann man sich die Aufgabe stellen: Es ist eine endliche Gruppe $\mathfrak{G}$ gegeben. Man soll mit ihr homomorphe lineare homogene Substitutionsgruppen mit Koeffizienten aus einem $GF[p^m]$ konstruieren. Vgl. Dickson, *Trans. Am. M. S.* **8**, 389 (1907), *Bull. Am. M. S.* (2) **13**, 477 (1907).

Wir führen nur folgendes Theorem von Blichfeldt (*Trans. Am. M. S.* **8**, 30 (1907)) an:

Ist $\mathfrak{G}$ eine irreduzible lineare homogene Substitutionsgruppe in n Variablen, so kann man stets unendlich viele derartige Primzahlen p finden, daß sich mittelst jeder von ihnen eine zu $\mathfrak{G}$ holoedrisch isomorphe irreduzible Gruppe $\mathfrak{G}_1$ linearer homogener Substitutionen in ebenfalls n Variablen bilden läßt, deren Koeffizienten ausschließlich ganze, mod p zu nehmende Zahlen sind.

Kapitel IV.

Algebraische Gleichungen.

Von *Alfred Loewy* in Freiburg i. Br.

§ 1. Historisches. Literatur.

Aufgabe der Algebra ist das Studium der ganzen rationalen Funktionen oder Polynome in einer oder mehreren Variablen (vgl. S. 31). Der Name „Algebra" stammt von dem Titel des mathematischen Lehrbuches *Aldschebr walmukâbala* des Arabers Muḥammed ibn Mûsâ Alchwarizmî (Anfang des 9. Jahrh.). Ihren Ausgangspunkt hat die Algebra in der Theorie der Gleichungen, und diese bilden auch ihren vornehmsten Gegenstand. Mit *Gleichungen des ersten Grades* beschäftigt sich bereits das altägyptische Rechenbuch des *Ahmes* (etwa 1700 v. Chr.). Die Auflösung der *Gleichungen zweiten Grades* war den griechischen Mathematikern bekannt. In geometrischer Form findet man sie in Euklids Elementen (um 300 v. Chr.) (Buch II, Satz 11, Buch VI, Satz 28 u. 29); zahlreichen quadratischen Gleichungen begegnet man in dem ältesten Denkmal griechischer algebraischer Wissenschaft, in Diophantus' Arithmetica (3. bis 4. Jahrh. n. Chr.). Die Auflösung der *Gleichungen dritten und vierten Grades* verdankt man italienischen Mathematikern des 16. Jahrhunderts: Scipione del Ferro, Tartaglia, Cardano und Ferrari (vgl. § 8).

Zu einem Herausgehen über ihre ersten Anfänge konnte die Algebra erst durch Schöpfung der *Buchstabenrechnung* gelangen. Vieta (*In artem analyticam isagoge* (1591)) hat zuerst die Buchstaben nicht nur wie bisher für die Unbekannten verwendet, sondern auch für Größen, die bestimmte gegebene Zahlenwerte besitzen. Vieta (*De aequationum recognitione et emendatione*, aus dem Nachlaß 1615 erschienen) hat auch den Zusammenhang zwischen den Wurzeln und Koeffizienten einer

Gleichung entdeckt; er beschränkte sich hierbei auf die positiven Lösungen, die er allein als Auflösungen zuließ. Girard (*Invention nouvelle en l'algèbre* (1629)) hat zuerst ausgesprochen: Jede Gleichung besitzt soviel Wurzeln, wie ihr Grad angibt; den Satz von Vieta hat er ohne Einschränkung, sogar beim Auftreten imaginärer oder mehrfacher Wurzeln. Er berechnet auch die Summen der vier ersten Potenzen der Gleichungswurzeln aus den Gleichungskoeffizienten. Descartes hat im dritten Buch seiner *Géométrie* (1637) (deutsche Ausgabe von Ludw. Schlesinger, Berlin 1894), welche seine berühmte Zeichenregel enthält und auch sonst reich an algebraischen Theoremen ist, den Satz: Wenn eine Gleichung mit der Unbekannten z die Wurzel α_1 besitzt, so ist ihre auf Null gebrachte linke Seite durch $z - \alpha_1$ ohne Rest teilbar. Auf Grund dieses Satzes erfordert der sogenannte *Fundamentalsatz oder Wurzelexistenzsatz* der Algebra nur den Nachweis, daß jede algebraische Gleichung wenigstens eine Wurzel besitzt; hieraus folgt, daß die auf Null gebrachte linke Seite jeder Gleichung n^{ten} Grades in n Faktoren ersten Grades zerfällt und daher, wenn man den Begriff der mehrfachen Wurzeln einführt, bei richtiger Zählung der Wurzeln genau n Wurzeln besitzt. Eine Regel zur Erkennung mehrfacher Wurzeln stammt von Hudde (*Huddenii epistolae* in Schootens Ausgabe von Cartesius' *Geometria* aus dem Jahre 1659, p. 433 u. 507). D'Alembert (*Recherches sur le calcul intégral, Histoire de l'acad. de Berlin, année 1746*) versuchte zuerst, einen Beweis für den Fundamentalsatz der Algebra zu liefern und hierdurch der Algebra eine einwandfreie Grundlage zu geben; daher bezeichnet man den Wurzelexistenzsatz auch als d'Alembert'sches Theorem. Den ersten auf strengerer Grundlage beruhenden Beweis des Fundamentalsatzes gab Gauß (*Ges. Werke* **3**, 1) in seiner Doktordissertation (1799) und hob hierbei die Mängel in den Beweisen seiner Vorgänger d'Alembert, Euler, de Foncenex und Lagrange hervor. In den Jahren 1815, 1816 und 1849 publizierte Gauß noch drei weitere Beweise für den Fundamentalsatz der Algebra (Gauß, *Ges. Werke* **3**, 31, 57, 71). Vgl. die vier Gaußschen Beweise für die Zerlegung ganzer algebraischer Funktionen in reelle Faktoren ersten oder zweiten Grades, herausgegeben von E. Netto, *Ostwalds Klassiker der exakten Wiss.*, Nr. 14.

Die Tatsache, daß sich die Wurzeln der Gleichungen der vier ersten Grade durch Radikale aus den Koeffizienten bilden lassen, veranlaßte die Mathematiker des 17. und 18. Jahrhunderts,

auch bei den Gleichungen höheren Grades nach einer *algebraischen Auflösung* zu suchen. Unter einer solchen versteht man eine Darstellung der Wurzeln aus den Koeffizienten mittelst einer endlichen Anzahl von Radikalen oder Wurzelzeichen, anders ausgedrückt: eine Rückführung der gegebenen Gleichung auf eine Kette reiner Gleichungen der Form $z^n = a$. Bei diesen Bemühungen gelangte Tschirnhaus, *Methodus auferendi omnes terminos intermedios ex data aequatione, Acta eruditorum* (1683), zu der heute nach seinem Namen sogenannten Tschirnhausentransformation. Lagranges *Réflexions sur la résolution algébrique des équations* (1770), *Œuvres* **3**, 205 (vgl. S. 168) verfolgen, wie es in der Einleitung heißt, den Zweck, a priori zu zeigen, warum die bei den Gleichungen der vier ersten Grade angewendeten Methoden bei den höheren Gleichungen versagen. Hierbei untersucht Lagrange die rationalen Funktionen der Gleichungswurzeln und ihre Werteänderungen bei Permutationen. Gleichzeitig und unabhängig von ihm haben sich auch Waring (*Miscellanea analytica* (1762), *Meditationes algebraicae* (1770), 3. ed. (1782)) und Vandermonde (vgl. unten, sowie Lagrange, *Œuvres* 8, 324) mit diesen Fragen beschäftigt. Von Lagrange entnahmen Ruffini und Abel die leitenden Ideen für den von ihnen erbrachten Beweis, daß die *allgemeine Gleichung von höherem als viertem Grade algebraisch nicht lösbar ist*, d. h. ihre Wurzeln sich nicht aus den Koeffizienten mit Hilfe von Radikalen ableiten lassen. Ruffini, *Teoria generale delle equazioni in cui si dimostra impossibile la soluzione algebraica delle equazioni generali di grado superiore al quarto, Bologna* (1799), vgl. Burkhardt, *Ztschr. für Math. u. Phys.* **37**, Suppl. S. 119 (1892), Abel, *Beweis der Unmöglichkeit, algebraische Gleichungen von höheren Graden als dem vierten allgemein aufzulösen, Journ. f. Math.* **1**, 65 (1826), *Œuvres par Sylow et Lie* (1881), **1**, 66, vgl. Pierpont, *Monatsh. f. Math.* **6**, 15 (1895).

Während die allgemeinen Gleichungen von höherem als viertem Grade nach dem Abel-Ruffinischen Theorem nicht algebraisch auflösbar sind, gibt es *spezielle Gleichungsgruppen*, deren Wurzeln sich *algebraisch* finden lassen. Vandermonde, *Mém. sur la résolution des équations, Histoire de l'acad. des sciences* (1771), deutsche Ausgabe, Berlin 1888, S. 63, hat als erster in der Gleichung $z^{11} = 1$ eine algebraisch auflösbare Gleichung kennen gelehrt und ihre Wurzeln mit Hilfe von Radikalen dargestellt. In Sectio 7 seiner *Disquitiones arithmeticae* (1801) (*Ges. Werke* **1**, 412) hat Gauß allgemein für die bei der *Kreis-*

teilung auftretenden Gleichungen, d. h. für die Gleichungen $z^n = 1$, die algebraische Auflösbarkeit nachgewiesen. (Vgl. § 12.) Eine große und zwar die einfachste und fundamentalste Klasse algebraisch auflösbarer Gleichungen lehrte Abel (*Journ. f. Math.* **4**, 131 (1829), *Œuvres par Sylow et Lie* **1**, 478, deutsche Ausg. mit Anm. von Loewy in *Ostwalds Klassikern der exakt. Wiss.*, Nr. 111) kennen. Diese Gleichungen werden nach Kronecker (*Monatsb. d. Berl. Akad.* (1853), 368, ebenda (1877), 846 und C. Jordan (*Traité*, p. 287)) als Abelsche Gleichungen bezeichnet. (Vgl. § 11.) Den größten Fortschritt in der Behandlung beliebiger algebraischer Gleichungen bedeuten die Arbeiten von Galois durch die Entdeckung der nach seinem Namen genannten Gruppe (vgl. S. 169). (Vgl. § 10.) Diese schreibt die für die Behandlung der betreffenden Gleichung einzuschlagenden Wege vor und sagt auch im besonderen aus, ob die Gleichung algebraisch auflösbar oder nicht auflösbar ist. C. Jordan nennt in höchster Bescheidenheit seinen fundamentalen *Traité des substitutions et des équations algébriques*, Paris 1870, einen bloßen Kommentar zu Galois' Werken.

Während die allgemeinen Gleichungen von höherem als viertem Grade nicht algebraisch lösbar sind, haben sie nach dem Fundamentalsatz der Algebra natürlich Lösungen. Der historische Weg, der dazu führte, die Wurzeln der *Gleichung fünften Grades* in ihrer Abhängigkeit von den Gleichungskoeffizienten darzustellen, war folgender: Die Transformation p^{ter} Ordnung (p ungerade Primzahl) der elliptischen Funktionen führte bereits Abel und Jacobi auf besondere algebraische Gleichungen $(p+1)^{\text{ten}}$ Grades, deren Wurzeln sich mittelst der Theorie der elliptischen Funktionen darstellen lassen. Die Galoissche Gruppe dieser Transformationsgleichungen ist nach Adjunktion einer Quadratwurzel die Modulargruppe der Ordnung $\frac{1}{2}p(p^2-1)$, also für $p \geqq 5$ eine einfache Gruppe. Die fraglichen Gleichungen sind daher infolge der Natur ihrer Galoisschen Gruppe nicht durch Radikale lösbar (vgl. die Darstellung bei H. Weber, *Algebra* **3**, 284, sowie bei Hölder, *Enzykl. d. math. Wiss.* **1**, 509). Für $p = 5$ besitzt die Modulargruppe eine Untergruppe des Index 5, daher muß sich für die fragliche Gleichung sechsten Grades eine rationale Resolvente (Hilfsgleichung) fünften Grades bilden lassen. Zu diesem gruppentheoretischen Ergebnis war bereits Galois (vgl. oben S. 215 letzte Zeile) gelangt, und Betti (1853) (*Opere mat.* **1**, 95) hat es bewiesen. Für die $p = 5$

entsprechende „Modulargleichung“ sechsten Grades, die zwischen der vierten Wurzel $\sqrt[4]{k}$ des Legendreschen Moduls und seinem transformierten Wert besteht, und deren Wurzeln sich explizit als elliptische Modulfunktionen darstellen lassen, hat Hermite (*C. R.* **46**, 508 (1858), *Œuvres* **2**, 5) wirklich eine Resolvente fünften Grades aufgestellt und ihre Wurzeln, die rationale Funktionen der Wurzeln der Modulargleichung sechsten Grades sind, angegeben. Diese Gleichung fünften Grades hat dieselbe Galoissche Gruppe wie die allgemeine Gleichung fünften Grades nach Adjunktion der Quadratwurzel ihrer Diskriminante, nämlich die alternierende Gruppe von fünf Symbolen. Mit der Bemerkung, daß sich nach den Untersuchungen des englischen Mathematikers Jerrard (1834) (die älteren des schwedischen Mathematikers Bring (1786) wurden erst im Jahre 1861 durch Hills Mitteilung an die schwedische Akademie der Vergessenheit entrissen) jede Gleichung fünften Grades ohne Verwendung anderer Irrationalitäten als Quadrat- und Kubikwurzeln in die obige Resolvente der Form $u^5 - u - A = 0$ mit nur einem Parameter A überführen läßt, hatte Hermite die erste Auflösung der Gleichung fünften Grades gewonnen. (Über die Methode von Bring und Hermite vgl. Klein, *Ikosaeder,* S 143 u. 244, über Bring vgl. Eneström, *Bibl. math.* (3) 8, 417 (1908)). Die bei Hermites oder ähnlichen Formeln stattfindende Benutzung elliptischer Modulfunktionen ist, wie F. Klein, *Ikosaeder,* S. 131, besonders *Journ. f. Math.* **129**, 154 (1905) hervorgehoben hat, ein Umweg, ebenso wie die Lösung der reinen Gleichung $z^n = a$ durch Logarithmen in der Form $z = e^{\frac{1}{n} i a}$.

Im gleichen Jahre 1858 wie Hermite publizierte auch Kronecker, *C. R.* **46**, 1150 (1858), *Monatsb. d. Berl. Akad.* (1861), 609 eine Auflösung der Gleichung fünften Grades. Er geht direkt von der allgemeinen Gleichung fünften Grades aus und macht ihre Auflösung von einer Gleichung zwölften Grades abhängig, deren Koeffizienten rationale Funktionen der Koeffizienten der Gleichung fünften Grades und der Quadratwurzel aus der Diskriminante sind. Die Resolvente ist zwar höheren Grades als die Ausgangsgleichung, aber insofern einfacher, weil zwischen ihren Wurzeln lineare Relationen bestehen, die für sie charakteristisch sind. Aus ihnen folgt, daß sich die Wurzeln der Gleichung zwölften Grades linear und homogen durch drei Parameter darstellen (Kiepert, *Journ. f. Math.* **87**, 115 (1879), Cayley, *Math. Ann.* **30**, 78 (1887), *Journ. f. Math.* **113**, 42 (1894), *Coll. math. papers* **12**, 493, **13**, 473). Führt man in der Gleichung zwölften Grades,

welche die Unbekannte nur quadratisch enthält, das Quadrat der Unbekannten als neue Unbekannte ein, so hat man eine Gleichung sechsten Grades. Auf solche Gleichungen ist zuerst Jacobi (*Journ. f. Math.* **3**, 308 (1828), *Ges. Werke* **1**, 261) in der Theorie der elliptischen Funktionen gekommen. Er hat nämlich gezeigt, daß, wenn p eine ungerade Primzahl und M der Multiplikator für eine Transformation p^{ter} Ordnung der elliptischen Funktionen ist, M Wurzel einer Gleichung $(p+1)^{\text{ten}}$ Grades ist, deren Koeffizienten rationale Funktionen des Moduls k sind. Die Quadratwurzeln der Lösungen einer solchen Gleichung $(p+1)^{\text{ten}}$ Grades lassen sich linear und homogen durch $\frac{p+1}{2}$ Parameter ausdrücken. Die eingehende Untersuchung und Verwertung der $p=5$ entsprechenden allgemeinen Jacobischen Gleichungen sechsten Grades für die Auflösung der Gleichung fünften Grades ist das Verdienst von Brioschi, dessen Arbeiten über diesen Gegenstand ebenfalls 1858 beginnen. Vgl. die Würdigung Brioschis durch Noether, *Math. Ann.* **50**, 483 (1898), sowie die zusammenfassende Darstellung von Brioschi, *Math. Ann.* **13**, 109 (1878). Über die Geschichte der Gleichung fünften Grades bis 1858 vgl. Pierpont, *Monatsh. f. Math.* **6**, 15 (1895).

Ein wesentlich neues Moment brachte Klein (vgl. seine zusammenfassende Darstellung in *Vorlesungen über das Ikosaeder*, Leipzig 1884) in die Theorie der Gleichung fünften Grades durch die Einführung der *Ikosaederirrationalität* (vgl. S. 240). Sie stellt sich neben die Radikale als neue und zwar ihrer Natur nach einfachste Transzendente; sie ist die Lösung einer Gleichung sechzigsten Grades mit nur einem Parameter, der Ikosaedergleichung, die in dem durch Adjunktion der fünften Einheitswurzeln erweiterten Körper[1]) ihre eigene Galoissche oder Normalgleichung ist, deren Galoissche Gruppe demnach die niedrigste einfachste Gruppe ist, die nicht Primzahlordnung besitzt. Die Auflösung der Gleichung fünften Grades besteht in zwei Schritten: der erste ist die Überführung einer beliebigen Gleichung fünften Grades in die Ikosaedergleichung und die Bestimmung der hierzu notwendigen Operationen, der zweite, transzendente Teil ist die Berechnung der Ikosaederirrationalität; diese wird durch die Verwendung hypergeometrischer Reihen ausgeführt, ähnlich wie sich die reinen

1) Die Erweiterung des Körpers ist deswegen erforderlich, weil die fünften Einheitswurzeln bei den Ikosaedersubstitutionen auftreten (vgl. S. 237).

Gleichungen $z^n = a$ durch binomische Reihen lösen lassen. Die Verwandschaft der Ikosaedergleichung mit Gleichungen sechsten Grades und deren Kenntnis aus der Theorie der elliptischen Funktionen erklärt, daß der historische Weg der Lösung der Gleichung fünften Grades über die Gleichungen sechsten Grades und die elliptischen Modulfunktionen führte. (Über die Behandlung der Gleichungen fünften und höheren Grades vgl. § 14.)

Den praktischen Rechner wird die *näherungsweise Berechnung der Wurzeln numerischer Gleichungen* besonders interessieren. Eine Vorläuferin der in Newtons *Analysis per aequationes* (1669) (M. Cantor, *Geschichte der Math.* **3**, 105) gelehrten „Newtonschen Näherungsmethode" findet man schon bei Vieta (1600) (Cantor **2**, 640). Rolles *Traité d'algèbre* (1690) und Newtons *Arithmetica universalis* (1707) geben Grenzen, zwischen denen die reellen Wurzeln einer algebraischen Gleichung liegen. Den größten Fortschritt in der Behandlung numerischer Gleichungen bedeuten die klassischen Arbeiten von Lagrange, Fourier und Sturm. Lagrange, *Traité de la résolution des équations numériques* (1798, nouv. éd. 1808), *Œuvres de Lagrange* **8**. Fourier, *Analyse des équations déterminées* (1831), deutsche Ausg. mit Anm. von A. Loewy in *Ostwalds Klassikern der exakt. Wiss.* Nr. 127. Ch. Sturm, *Mémoire sur la résolution des équations numériques, Mém. prés. par div. sav. à l'acad. des sc. de France* (1835), deutsche Ausg. mit Anm. von A. Loewy in *Ostwalds Klassikern der exakt. Wiss.* Nr. 143. Sturms Aufsatz hat zum erstenmal die genaue Anzahl reeller Wurzeln einer numerischen Gleichung mit reellen Koeffizienten zwischen zwei gegebenen reellen Zahlen bestimmen gelehrt, während die früheren Kriterien nur Grenzen für jene Zahl lieferten.

Die voraufgegangenen Zeilen bezwecken nicht, eine Geschichte der Algebra zu geben; sie wollen nur einige orientierende Bemerkungen machen. Über die ältere Geschichte der Algebra vgl. man M. Cantor, *Vorl. über Geschichte der Math.* Bd. 1—4, die systematische Darstellung bei Tropfke, *Geschichte der Elementarmathematik,* Leipzig 1902, **1**, 123, L. Matthiessen, *Grundzüge der antiken und modernen Algebra der litteralen Gleichungen*, Leipzig 1878 (am Schluß reichhaltiges Literaturverzeichnis).

Von Lehrbüchern nennen wir:

H. Weber, *Lehrbuch der Algebra,* Bd. **1**, **2** und **3**, Braunschweig, 2. Aufl., 1898, 1899 u. 1908. E. Netto, *Vorlesungen über Algebra,* Bd. **1** u. **2**, Leipzig 1896 u. 1900. J. A. Serret,

Handbuch der höheren Algebra, deutsch von Wertheim, Bd. **1** u. **2**, Leipzig, 2. Aufl., 1878 u. 1879. G. Bauer, *Vorlesungen über Algebra,* Leipzig 1903. J. Petersen, *Theorie der algebraischen Gleichungen,* Kopenhagen 1878. Runge, *Praxis der Gleichungen*, Samml. Schubert **14**, Leipzig 1900. Cesàro, *Elementares Lehrbuch der algebraischen Analysis und der Infinitesimalrechnung,* deutsch herausg. von Kowalewski, Leipzig 1904, S. 368—485. A. Loewy, *Kurzgefaßtes Lehrbuch der Algebra,* erscheint bei Veit & Co., Leipzig 1910. W. S. Burnside and A. Panton, *Theory of equations,* 4. Aufl., London 1900. M. Bôcher, *Introduction to higher algebra*, New York 1908, deutsche Ausg. erscheint 1909 bei Teubner, Leipzig. F. Cajori, *Introduction to the modern theory of equations*, New York 1904. L. E. Dickson, *Introduction to the theory of algebraic equations,* New York 1903. Borel et Drach, *Introduction à l'étude de la théorie des nombres et de l'algèbre supérieure*, Paris 1895. J. Tannery, *Leçons d'algèbre et d'analyse,* Paris 1906, 2 vol. Vogt, *Leçons sur la résolution algébrique des équations,* Paris 1895. Bianchi, *Lezioni sulla teoria dei gruppi di sostituzioni e delle equazioni algebriche secondo Galois,* Pisa 1900. Capelli, *Istituzioni di analisi algebrica*, 3. ed., Napoli 1902.

Wir verweisen noch auf die Artikel in der *Encyklopädie der math. Wiss.*, Bd. I, „Rationale Funktionen einer Veränderlichen; ihre Nullstellen" von Netto, S. 227, „Rationale Funktionen mehrerer Veränderlichen" von Netto, S. 2[illegible], „Separation und Approximation der Wurzeln" von Runge, S. 404, „Rationale Funktionen der Wurzeln; symmetrische und Affektfunktionen" von Vahlen, S. 449, „Galoissche Theorie mit Anwendungen" von Hölder, S. 480. In der *Encyclopédie des sciences math.* liegen in französischer Bearbeitung nach den Artikeln von Netto vor: *Les fonctions rationelles* von Le Vavasseur.

§ 2. Rationale Funktionen. Teilbarkeitsgesetze.

Bedeuten $a_0, a_1, a_2, \ldots, a_n$ beliebig gegebene reelle oder komplexe Größen und z eine Veränderliche, so heißt, wenn $a_0 \neq 0$ ist, der Ausdruck $a_0 z^n + a_1 z^{n-1} + a_2 z^{n-2} + \cdots + a_n$ eine *ganze rationale Funktion n^{ten} Grades* oder *ein Polynom n^{ten} Grades* in z. Eine solche ganze rationale Funktion von z bezeichnet man abgekürzt mit $f(z)$ oder $g(z)$ oder $P(z)$ oder $\varphi(z)$; hat man mehrere ganze Funktionen von z, so fügt man auch Indices bei und schreibt $f_1(z)$, $f_2(z)$ usw. *Das Produkt*

$f(z) \cdot g(z)$ zweier ganzer Funktionen $f(z)$ vom m^{ten} und $g(z)$ vom n^{ten} Grade ist ein Polynom vom Grade $m + n$.

Der Quotient $\frac{f(z)}{g(z)}$ zweier ganzer Funktionen heißt eine *gebrochene rationale Funktion.* Ist $f(z)$ von niedrigerem Grade als $g(z)$, so heißt $\frac{f(z)}{g(z)}$ *echt gebrochen. Ist $f(z)$ eine ganze Funktion vom m^{ten} Grad und $g(z)$ eine solche vom n^{ten} Grad und $m \geqq n$,* so gibt es zwei eindeutig bestimmte ganze Funktionen $G_1(z)$ vom Grad $m - n$[1]) und $R_1(z)$ vom $n - 1^{ten}$ oder niedrigeren Grade, daß $f(z) = g(z)G_1(z) + R_1(z)$ wird. Mithin ist $\frac{f(z)}{g(z)} = G_1(z) + \frac{R_1(z)}{g(z)}$; $G_1(z)$ heißt der Quotient, $R_1(z)$ der Rest der Division. Ist der Rest $R_1(z)$ gleich Null, so heißt $f(z)$ durch $g(z)$ ohne Rest teilbar; $\frac{f(z)}{g(z)}$ ist in diesem Falle, wie man sagt, nur scheinbar gebrochen. $g(z)$ heißt ein *Teiler* oder *Divisor* von $f(z)$. *Jede ganze Funktion ist durch jede Konstante und durch sich selbst teilbar.*

Ist α irgendeine Konstante, so ist $f(z) - f(\alpha)$ stets durch $z - \alpha$ ohne Rest teilbar.

Ist $f(z)$ eine ganze Funktion m^{ten} Grades und $g(z)$ vom n^{ten} Grade $m \geqq n$, so kann man durch fortgesetzte Division folgende Gleichungskette aufstellen:

$$\begin{aligned} f(z) &= g(z)G_1(z) + R_1(z), \\ g(z) &= R_1(z)G_2(z) + R_2(z), \\ &\vdots \qquad\qquad \vdots \\ R_{\lambda-2}(z) &= R_{\lambda-1}(z)G_\lambda(z) + R_\lambda. \end{aligned}$$

Die sukzessiv auftretenden Reste $R_1, R_2, \ldots, R_\lambda$ sind ganze Funktionen, deren Grade abnehmen. *Das Verfahren, unter dem Namen „Euclidscher Algorithmus" bekannt, läßt sich so weit treiben, bis man zu einem Rest R_λ kommt, der eine Konstante ist.* Ist R_λ eine von Null verschiedene Konstante, so sind die zwei Funktionen $f(z)$ und $g(z)$ *teilerfremd* oder *relativ prim,* d. h. es gibt außer Konstanten keine ganze Funktion, die sowohl $f(z)$ als auch $g(z)$ ohne Rest teilt. Man sagt auch: $f(z)$ und $g(z)$ haben keinen gemeinsamen Divisor. Ist R_λ gleich Null, so haben $f(z)$ und $g(z)$ die ganze Funktion $R_{\lambda-1}(z)$ zum gemeinsamen Teiler. *$R_{\lambda-1}(z)$ ist der größte gemeinsame Teiler von*

---

1) Eine Konstante ist eine ganze Funktion vom Grad Null.

$f(z)$ und $g(z)$, d. h. jeder gemeinsame Teiler von $f(z)$ und $g(z)$ ist Teiler von $R_{\lambda-1}$. Der größte gemeinsame Teiler zweier ganzer Funktionen ist eine bis auf eine multiplikative Konstante eindeutig bestimmte ganze Funktion. Der Euclidsche Algorithmus und die Bestimmung des größten gemeinsamen Teilers zweier ganzer rationaler Funktionen läßt sich mittelst ausschließlich rationaler Rechenoperationen ausführen.

Die charakteristische Bedingung dafür, daß zwei ganze Funktionen $f(z)$ und $g(z)$ teilerfremd sind, besteht in der Existenz zweier ganzer Funktionen $P(z)$ und $Q(z)$ von der Beschaffenheit, daß $Pf + Qg = 1$ ist. Ist $f(z)$ eine ganze Funktion vom m^{ten} Grad und $g(z)$ vom n^{ten} Grad, so lassen sich, wenn die zwei Funktionen f und g teilerfremd sind, auf eine einzige Art zwei Funktionen P_0 höchstens vom Grad $n-1$ und Q_0 höchstens vom Grad $m-1$ bestimmen, daß $P_0 f + Q_0 g = 1$ ist. Mittelst P_0 und Q_0 drücken sich P und Q in der Form $P_0 + gS$ und $Q_0 - fS$ aus, wobei S irgendeine ganze Funktion von z bedeutet.

Sind $f(z)$ und $g(z)$ teilerfremde ganze Funktionen und $U(z)$ eine derartige ganze Funktion von z, daß das Produkt $U(z)f(z)$ durch $g(z)$ ohne Rest teilbar ist, so ist $U(z)$ durch $g(z)$ teilbar.

Die charakteristische Bedingung dafür, daß zwei ganze Funktionen $f(z)$ vom Grad m und $g(z)$ vom Grad n einen gemeinsamen Teiler haben, besteht in der Existenz zweier ganzer Funktionen $G(z)$ höchstens $n-1^{ten}$ Grades und $F(z)$ höchstens $m-1^{ten}$ Grades von der Beschaffenheit, daß $fG + gF = 0$ ist. Ist G vom Grad $n-k$ und F demnach vom Grad $m-k$, so haben f und g einen gemeinsamen Teiler, der mindestens vom Grad k ist.

Sind $f_1(z), f_2(z), \ldots, f_r(z)$ irgendwelche ganze Funktionen von z, so heißt die ganze Funktion niedrigsten Grades, welche durch jede der Funktionen $f_1, f_2, \ldots, f_r$ ohne Rest teilbar ist, *das kleinste gemeinsame Vielfache von $f_1, f_2, \ldots, f_r$*.

Man kann auch *ganze rationale Funktionen* oder *Polynome mehrerer Veränderlichen* betrachten. Hierunter versteht man, wenn $z_1, z_2, \ldots, z_r$ Variable bedeuten, einen Ausdruck, der aus einer endlichen Anzahl von Summanden der Form

$$a_{\alpha_1 \alpha_2 \ldots \alpha_r} z_1^{\alpha_1} z_2^{\alpha_2} \ldots z_r^{\alpha_r}$$

besteht; $\alpha_1, \alpha_2, \ldots, \alpha_r$ bedeuten ganze positive *Zahlen* einschließlich der Null, $a_{\alpha_1 \alpha_2 \ldots \alpha_r}$ beliebige Konstanten. Unter dem *Grad der ganzen Funktion* der r Variablen versteht man den Maximalwert, den $\alpha_1 + \alpha_2 + \cdots + \alpha_r$ bei allen Gliedern mit

$a_{\alpha_1 \alpha_2 \ldots \alpha_r} \neq 0$ annimmt. *Die Anzahl der Glieder einer ganzen rationalen Funktion n^{ten} Grades in r Variablen ist* $\leqq \binom{n+r}{r}$. *Das Produkt zweier ganzer Funktionen ist wieder eine ganze Funktion, deren Grad gleich der Summe der Grade der Faktoren ist.*

Eine ganze rationale Funktion von einer oder mehreren Veränderlichen heißt *unzerlegbar* oder *absolut irreduzibel*, wenn sie nicht als das Produkt zweier ganzer rationaler Funktionen dargestellt werden kann, von denen jede von niedrigerem Grad als die ursprüngliche Funktion ist. *Die einzigen unzerlegbaren ganzen Funktionen in einer Veränderlichen sind die Funktionen ersten Grades, hingegen gibt es unzerlegbare Funktionen mehrerer Variablen von höherem als erstem Grad,* z. B. ist $z_1^2 - z_2^3$ eine unzerlegbare Funktion.

Wenn eine ganze rationale Funktion der r Variablen $z_1, z_2, \ldots, z_r$ in zwei Faktoren zerlegbar ist, die in bezug auf die eine Variable, etwa z_1, ganz, in bezug auf die anderen Variablen $z_2, z_3, \ldots, z_r$ wenigstens rational sind, so ist sie auch als Produkt zweier Funktionen darstellbar, die in allen r Variablen ganz und rational sind.

Eine ganze rationale Funktion kann, wenn man von konstanten Faktoren absieht, nur auf eine Art in ein Produkt unzerlegbarer Faktoren zerlegt werden.

Über ganze rationale Funktionen mehrerer Variablen vgl. Weber, *Algebra* **1**, 71, Netto, *Algebra* **2**, 1, Méray, *Nouv. Ann. de math.* (4) **7**, 193 (1907). Über Teilbarkeitsgesetze der ganzen rationalen Funktionen bei Zugrundelegung eines Körpers vgl. § 9, wo man auch weitere Literatur findet.

§ 3. Fundamentalsatz der Algebra. Einfache und mehrfache Wurzeln.

Hat man eine ganze rationale Funktion $g(z)$ einer Veränderlichen z, so kann man jene besonderen Werte von z aufsuchen, die, an die Stelle von z gesetzt, das Polynom $g(z)$ identisch zu Null machen. Sie heißen die *Wurzeln* oder *Lösungen der Gleichung* $g(z) = 0$. Der Grad des Polynoms $g(z)$ heißt der *Grad der Gleichung*.

Als notwendig und hinreichend, damit eine Größe α_1 Wurzel der Gleichung $g(z) = 0$ ist, erweist sich die Teilbarkeit von $g(z)$ durch $z - \alpha_1$ (Descartes, *Géométrie* (1637), vgl. oben S. 251).

Fundamentalsatz der Algebra: Jede algebraische Gleichung mit beliebigen Koeffizienten hat wenigstens eine Wurzel. Über

den Fundamentalsatz vgl. oben S. 251, ferner die historisch-kritische Studie von G. Loria, *Rivista di mat.*, Bd. 1, *Bibl. math.*, N. F. 5, 99 (1891), 7, 47 (1893), sowie die eingehende Besprechung der verschiedenen Beweise in dem Artikel „Les fonctions rationelles" (Netto-Le Vavasseur) der *Encycl. des sc. math.* 1, vol. 2, 189.

Aus dem Fundamentalsatz der Algebra folgt: *Jede ganze rationale Funktion* $g(z) = a_0 z^n + a_1 z^{n-1} + \cdots + a_n$ $(a_0 \neq 0)$ *läßt sich in ein Produkt* $a_0(z-\alpha_1)(z-\alpha_2)\ldots(z-\alpha_n)$ *zerlegen. Die Größen* $\alpha_1, \alpha_2, \ldots, \alpha_n$ *und auch nur sie allein sind die Wurzeln der Gleichung* $g(z) = 0$.

Die Größen $\alpha_1, \alpha_2, \ldots, \alpha_n$ brauchen nicht sämtlich voneinander verschieden zu sein; um den Satz aussprechen zu können, daß eine jede Gleichung n^{ten} Grades soviel Wurzeln besitzt, wie ihr Grad angibt, ist der *Begriff der mehrfachen Wurzeln* einzuführen. Tritt unter den Größen $\alpha_1, \alpha_2, \ldots, \alpha_n$ die Größe α_1 genau r_1-fach, α_2 genau r_2-fach usw., α_k genau r_k-fach auf und ist $r_1 + r_2 + \cdots + r_k = n$, so heißt α_1 eine *r_1-fache*, α_2 eine *r_2-fache*, $\ldots$, α_k eine *r_k-fache Wurzel der Gleichung* $g(z) = 0$.

Ist die Größe α eine r-fache Wurzel der Gleichung $g(z) = 0$, so ist $g(z)$ durch $(z-\alpha)^r$ und keine höhere Potenz von $z - \alpha$ als die r-te ohne Rest teilbar.

Die Größe α ist dann und nur dann eine r-fache Wurzel von $g(z) = 0$, wenn α die ganze rationale Funktion $g(z)$ samt ihren ersten $r-1$ Abgeleiteten, hingegen nicht mehr $g^{(r)}(z)$ annulliert (Huddesche Regel vgl. oben S. 251).

Die *erste Abgeleitete der ganzen Funktion*

$$g(z) = a_0 z^n + a_1 z^{n-1} + a_2 z^{n-2} + \cdots + a_n$$

lautet

$$g'(z) = n a_0 z^{n-1} + (n-1) a_1 z^{n-2} + (n-2) a_2 z^{n-3} + \cdots + 2 a_{n-2} z + a_{n-1}.$$

(Vgl. Differentialrechnung.) Die k^{te} Abgeleitete einer Funktion $g(z)$ ist die erste Abgeleitete der $(k-1)^{\text{ten}}$ Abgeleiteten.

Eine Gleichung $g(z) = 0$ hat dann und nur dann keine mehrfachen Wurzeln, wenn die Funktion $g(z)$ und ihre erste Abgeleitete $g'(z)$ teilerfremd sind. Ist $D_1(z)$ der größte gemeinsame Teiler von $g(z)$ und $g'(z)$, so ist $\frac{g(z)}{D_1(z)} = G_1(z)$ eine ganze Funktion. Die Gleichung $G_1(z) = 0$ hat nur einfache Wurzeln und wird durch sämtliche Wurzeln von $g(z) = 0$ befriedigt.

Der größte gemeinsame Teiler der ganzen Funktion $D_i(z)$ und ihrer ersten Abgeleiteten $D_i'(z)$ sei mit $D_{i+1}(z)$ bezeichnet. Man bilde aus $D_1(z)$ die Funktion $D_2(z)$, dann $D_3(z)$ usw. *Die Gleichung $g(z) = 0$ hat dann und nur dann i-fache, aber keine $(i+1)$-fachen Wurzeln, wenn D_i eine Konstante wird und zwar die erste auftretende Konstante ist.* Die Gleichung $D_i(z) = 0$ wird nämlich nur durch mindestens $(i+1)$-fache Wurzeln der Gleichung $g(z) = 0$ befriedigt, und jede tritt in einer um i niedrigeren Multiplizität als bei der Gleichung $g(z) = 0$ auf.

Bildet man die ganzen, nur scheinbar gebrochenen Funktionen

$$\frac{g(z)}{D_1(z)} = G_1(z), \quad \frac{D_1(z)}{D_2(z)} = G_2(z), \quad \frac{D_2(z)}{D_3(z)} = G_3(z), \ldots,$$

so haben sämtliche Gleichungen $G_1(z) = 0$, $G_2(z) = 0$, $G_3(z) = 0, \ldots$ nur einfache Wurzeln und keine anderen Wurzeln als solche der Gleichung $g(z) = 0$, und zwar wird $G_1(z) = 0$ durch alle Wurzeln der Gleichung $g(z) = 0$ erfüllt, $G_2(z) = 0$ durch alle mindestens zweifachen, $G_3(z) = 0$ durch alle mindestens dreifachen usw.

Hat eine Gleichung $g(z) = 0$ mit reellen Koeffizienten die imaginäre Größe $\alpha + i\beta$ zur r-fachen Wurzel, so ist auch die konjugiert imaginäre Größe $\alpha - i\beta$ eine r-fache Wurzel der Gleichung $g(z) = 0$.

§ 4. Partialbruchzerlegung einer gebrochenen rationalen Funktion.

Jede nicht echt gebrochene Funktion $\frac{f(z)}{g(z)}$ läßt sich als Summe einer ganzen Funktion $G_1(z)$ und einer echt gebrochenen Funktion $\frac{R(z)}{g(z)}$ darstellen (vgl. S. 258).

Ist $\frac{R(z)}{g(z)}$ echt gebrochen, also $R(z)$ eine ganze Funktion niedrigeren Grades als $g(z)$, und hat $g(z)$ die Form

$$g(z) = (z-\alpha_1)^{r_1}(z-\alpha_2)^{r_2}\ldots(z-\alpha_k)^{r_k},$$

so daß $\alpha_1, \alpha_2, \ldots, \alpha_k$ bezw. r_1-, r_2-, $\ldots$, r_k-fache Wurzeln von $g(z) = 0$ bedeuten, so wird

$$\begin{aligned}\frac{R(z)}{g(z)} = &\frac{A_1}{(z-\alpha_1)^{r_1}} + \frac{A_2}{(z-\alpha_1)^{r_1-1}} + \cdots + \frac{A_{r_1}}{z-\alpha_1} + \\ &\frac{B_1}{(z-\alpha_2)^{r_2}} + \frac{B_2}{(z-\alpha_2)^{r_2-1}} + \cdots + \frac{B_{r_2}}{z-\alpha_2} + \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ &\frac{K_1}{(z-\alpha_k)^{r_k}} + \frac{K_2}{(z-\alpha_k)^{r_k-1}} + \cdots + \frac{K_{r_k}}{z-\alpha_k};\end{aligned}$$

hierbei sind $A_1, A_2, \ldots, A_{r_1}, B_1, B_2, \ldots, B_{r_2}, K_1, K_2, \ldots, K_{r_k}$ *eindeutig bestimmte Konstanten. Die Größen A werden durch die folgenden Rekursionsformeln gefunden, bei denen die in Klammern stehenden oberen Indices Abgeleitete bedeuten* (vgl. Differentialrechnung):

$$A_1 \frac{g^{(r_1)}(\alpha_1)}{r_1!} = R(\alpha_1),$$

$$A_2 \frac{g^{(r_1)}(\alpha_1)}{r_1!} + A_1 \frac{g^{(r_1+1)}(\alpha_1)}{(r_1+1)!} = \frac{R'(\alpha_1)}{1},$$

$$A_3 \frac{g^{(r_1)}(\alpha_1)}{r_1!} + A_2 \frac{g^{(r_1+1)}(\alpha_1)}{(r_1+1)!} + A_1 \frac{g^{(r_1+2)}(\alpha_1)}{(r_1+2)!} = \frac{R''(\alpha_1)}{1 \cdot 2},$$

. . . . . .

Da α_1 *eine* r_1*-fache Wurzel von* $g(z) = 0$ *ist, wird* $g^{(r_1)}(\alpha_1) \neq 0$. *Analoge Formeln gelten für die Größen* $B_1, B_2, \ldots, B_{r_2}$ *usw.* Eine derartige Zerlegung bezeichnet man als eine Zerlegung einer *gebrochenen rationalen Funktion in Partialbrüche;* sie ist für die Integralrechnung besonders wichtig. Leibniz und Johann Bernoulli haben sich bereits 1702 und 1703 mit ihr zuerst für $r_1 = r_2 = \cdots = r_k = 1$, dann für beliebige $r_1, r_2, \ldots, r_k$ beschäftigt (vgl. M. Cantors *Vorl. über Geschichte der Math.* **3**, 272).

Für $r_1 = r_2 = \cdots = r_k = 1$ erhält man die einfachen Formeln $A_1 = \frac{R(\alpha_1)}{g'(\alpha_1)}$, $B_1 = \frac{R(\alpha_2)}{g'(\alpha_2)}, \ldots, K_1 = \frac{R(\alpha_k)}{g'(\alpha_k)}$. Hieraus ergibt sich die sogenannte *Lagrangesche Interpolationsformel* (Lagrange, *Leçons élémentaires* (1795), *Œuvres* **7**, 285, früher bei Waring, *Phil. Trans.* **69** (1779), vgl. Braunmühl, *Bibl. math.* (3) **2**, 95 (1901)): *Eine ganze rationale Funktion* $R(z)$*, die höchstens vom Grad* $k-1$ *ist und für die k untereinander verschiedenen Werte* $\alpha_1, \alpha_2, \ldots, \alpha_k$ *die k vorgegebenen Werte* $u_1 = R(\alpha_1)$, $u_2 = R(\alpha_2), \ldots, u_k = R(\alpha_k)$ *annimmt, hat den Wert:*

$$R(z) = u_1 \frac{g(z)}{(z-\alpha_1)g'(\alpha_1)} + u_2 \frac{g(z)}{(z-\alpha_2)g'(\alpha_2)} + \cdots + u_k \frac{g(z)}{(z-\alpha_k)g'(\alpha_k)};$$

hierbei ist

$$g(z) = (z-\alpha_1)(z-\alpha_2)\cdots(z-\alpha_k),$$

$$g'(\alpha_i) = (\alpha_i-\alpha_1)(\alpha_i-\alpha_2)\cdots(\alpha_i-\alpha_{i-1})(\alpha_i-\alpha_{i+1})\cdots(\alpha_i-\alpha_k)$$
$$(i = 1, 2, \ldots, k).$$

Es gibt nur eine solche Funktion $R(z)$.

In innigstem Zusammenhang mit der Lagrangeschen Interpolationsformel stehen die Eulerschen Formeln (Euler,

Institutiones calculi integralis (1769) vol. 2, § 1169 und 1170): *Hat die Gleichung* $g(z) = 0$ *lauter verschiedene Wurzeln* $\alpha_1, \alpha_2, \ldots, \alpha_n$, *so ist:*

$$\frac{{\alpha_1}^m}{g'(\alpha_1)} + \frac{{\alpha_2}^m}{g'(\alpha_2)} + \cdots + \frac{{\alpha_n}^m}{g'(\alpha_n)} = 0 \qquad (m = 0, 1, 2, \ldots, n-2),$$

$$\frac{{\alpha_1}^{n-1}}{g'(\alpha_1)} + \frac{{\alpha_2}^{n-1}}{g'(\alpha_2)} + \cdots + \frac{{\alpha_n}^{n-1}}{g'(\alpha_n)} = \frac{1}{g_0},$$

wenn g_0 *der Koeffizient der höchsten Potenz von* z *in* $g(z)$ *ist.* Hieraus folgt: *Für jede ganze rationale Funktion* $R(z)$, *deren Grad um wenigstens zwei Einheiten niedriger als der von* $g(z)$ *ist, besteht, wenn* $g(z) = 0$ *lauter verschiedene Wurzeln* $\alpha_1, \alpha_2, \ldots, \alpha_n$ *besitzt, die Gleichung:*

$$\frac{R(\alpha_1)}{g'(\alpha_1)} + \frac{R(\alpha_2)}{g'(\alpha_2)} + \cdots + \frac{R(\alpha_n)}{g'(\alpha_n)} = 0.$$

Sind $g_1(z), g_2(z), \ldots, g_k(z)$ *ganze Funktionen von* z, *von denen je zwei relativ prim sind, ist* $\frac{R(z)}{g(z)}$ *eine echt gebrochene Funktion und* $g(z) = g_1(z)^{r_1} g_2(z)^{r_2} \ldots g_k(z)^{r_k}$, *so besteht die Gleichung:*

$$\begin{aligned}\frac{R(z)}{g(z)} = {} & \frac{A_1}{g_1(z)^{r_1}} + \frac{A_2}{g_1(z)^{r_1-1}} + \cdots + \frac{A_{r_1}}{g_1(z)} + \\ & + \frac{B_1}{g_2(z)^{r_2}} + \frac{B_2}{g_2(z)^{r_2-1}} + \cdots + \frac{B_{r_2}}{g_2(z)} + \\ & \vdots \\ & + \frac{K_1}{g_k(z)^{r_k}} + \frac{K_2}{g_k(z)^{r_k-1}} + \cdots + \frac{K_{r_k}}{g_k(z)};\end{aligned}$$

hierbei sind $A_1, A_2, \ldots, A_{r_1}$ *eindeutig bestimmte ganze Funktionen von* z, *deren Grad niedriger als* $g_1(z)$ *ist,* $B_1, B_2, \ldots, B_{r_2}$ *eindeutig bestimmte ganze Funktionen von* z, *deren Grad niedriger als* $g_2(z)$ *ist, usw., schließlich* $K_1, K_2, \ldots, K_{r_k}$ *eindeutig bestimmte ganze Funktionen, deren Grad niedriger als* $g_r(z)$ *ist.*

Der erste Satz dieses Paragraphen ist nur ein Spezialfall der zuletzt angegebenen Formel. Ferner ergibt sich aus ihr: *Ist* $g(z) = (az^2 + bz + c)^{r_1} g_2(z)$ *und sind* $g_2(z)$ *und* $az^2 + bz + c$ *relativ prim, so hat man, wenn* $R(z)$ *von niedrigerem Grad als* $g(z)$ *ist, die Zerlegung:*

$$\frac{R(z)}{g(z)} = \frac{A_1 z + B_1}{(az^2+bz+c)^{r_1}} + \frac{A_2 z + B_2}{(az^2+bz+c)^{r_1-1}} + \cdots + \frac{A_{r_1} z + B_{r_1}}{az^2+bz+c} + \frac{F(z)}{g_2(z)};$$

hierbei ist $F(z)$ von niedrigerem Grad als $g_2(z)$, die Größen A und B bedeuten Konstante.

Jede ganze Funktion $g(z)$ mit reellen Koeffizienten läßt sich in ein Produkt von reellen Faktoren ersten Grades $z - \alpha$ und zweiten Grades $az^2 + bz + c$ zerlegen. Hieraus folgt auf Grund des voraufgegangenen Satzes: *Haben in der echt gebrochenen Funktion $\frac{R(z)}{g(z)}$ die ganzen Funktionen $R(z)$ und $g(z)$ reelle Koeffizienten, so läßt sich $\frac{R(z)}{g(z)}$ als eine Summe elementarer Brüche der Form $\frac{A}{(z-\alpha)^\lambda}$ und $\frac{Bz+C}{(az^2+bc+c)^\mu}$ darstellen, wobei A, B, C, α, a, b, c lauter reelle Größen sind.* Diese Zerlegung ist für die Integralrechnung wichtig.

§ 5. Symmetrische Funktionen der Gleichungswurzeln.

Hat man die Gleichung $a_0 z^n + a_1 z^{n-1} + \cdots + a_n = 0$ mit den n Wurzeln $\alpha_1, \alpha_2, \ldots, \alpha_n$, so drücken sich die symmetrischen Grundfunktionen der Wurzeln durch die Gleichungskoeffizienten in der Form aus (vgl. für das Folgende S. 219):

$$S_1 = \alpha_1 + \alpha_2 + \cdots + \alpha_n = -\frac{a_1}{a_0},$$

$$S_2 = \sum \alpha_1 \alpha_2 = +\frac{a_2}{a_0},$$

$$S_3 = \sum \alpha_1 \alpha_2 \alpha_3 = -\frac{a_3}{a_0},$$

$$\vdots$$

$$S_n = \alpha_1 \cdot \alpha_2 \ldots \alpha_n = (-1)^n \frac{a_n}{a_0}.$$

(Vieta (1615), Girard (1629), vgl. S. 250.)

Jede ganze symmetrische Funktion der Gleichungswurzeln $\alpha_1, \alpha_2, \ldots, \alpha_n$ mit ganzzahligen Koeffizienten ist auf eine und auch nur auf eine Weise als ganze ganzzahlige Funktion der Quotienten $\frac{a_1}{a_0}, \frac{a_2}{a_0}, \ldots, \frac{a_n}{a_0}$ darstellbar. Jede typische symmetrische Funktion $(\nu_1, \nu_2, \ldots, \nu_n)$ (vgl. S. 218) *der Gleichungswurzeln $\alpha_1, \alpha_2, \ldots, \alpha_n$ wird durch Multiplikation mit $a_0^{\nu_1}$, je-*

doch keiner niedrigeren als der ν_1^{ten} Potenz von a_0, eine ganze homogene Funktion der Größen $a_0, a_1, \ldots, a_n$, also eine Summe von Gliedern der Form $a_0^{\lambda_0} a_1^{\lambda_1} a_2^{\lambda_2} \ldots a_n^{\lambda_n}$, mit ganzzahligen Zahlenkoeffizienten vom Grad ν_1; es ist also $\lambda_0 + \lambda_1 + \cdots + \lambda_n = \nu_1$ (Folge des Waringschen Resultates auf S. 219, vgl. Sylvester, *Coll. math. papers* **1**, 595).

Hat man ein Glied $Z_{\lambda_0 \lambda_1 \ldots \lambda_n} a_0^{\lambda_0} a_1^{\lambda_1} a_2^{\lambda_2} \ldots a_n^{\lambda_n}$, wobei $Z_{\lambda_0 \lambda_1 \ldots \lambda_n}$ einen bloßen Zahlenkoeffizienten bedeutet, so bezeichnet man $\lambda_1 + 2\lambda_2 + 3\lambda_3 + \cdots + n\lambda_n$ als das *Gewicht des Gliedes*. Hat man eine Funktion, die aus einer Summe von Gliedern der Form $Z_{\lambda_0 \lambda_1 \ldots \lambda_n} a_0^{\lambda_0} a_1^{\lambda_1} a_2^{\lambda_2} \ldots a_n^{\lambda_n}$ besteht, und haben alle Glieder das nämliche Gewicht $\lambda_1 + 2\lambda_2 + 3\lambda_3 + \cdots + n\lambda_n$, so heißt der Ausdruck *isobar*.

Jede typische symmetrische Funktion $(\nu_1, \nu_2, \ldots, \nu_n)$ der Gleichungswurzeln ist nach Multiplikation mit $a_0^{\nu_1}$ als ganze homogene isobare Funktion der Gleichungskoeffizienten $a_0, a_1, a_2, \ldots, a_n$ vom Gewicht $\nu_1 + \nu_2 + \cdots + \nu_n$ ausdrückbar (Cayley, *Coll. math. papers* **2**, 418).

Die ganzen Potenzen $s_\nu = (\nu, 0, 0, \ldots, 0) = \alpha_1^\nu + \alpha_2^\nu + \cdots + \alpha_n^\nu$ der Gleichungswurzeln lassen sich in rekurrenter Weise durch die Gleichungskoeffizienten mittelst der Newtonschen Formeln (Newton, *Arithmetica universalis* (1707)) *bestimmen:*

$$\begin{aligned}
a_0 s_1 + a_1 &= 0,\\
a_0 s_2 + a_1 s_1 + 2a_2 &= 0,\\
a_0 s_3 + a_1 s_2 + a_2 s_1 + 3a_3 &= 0,\\
&\cdots\cdots\cdots\cdots\cdots\\
a_0 s_n + a_1 s_{n-1} + a_2 s_{n-2} + \cdots + n a_n &= 0,\\
a_0 s_{n+k} + a_1 s_{n+k-1} + a_2 s_{n+k-2} + \cdots + a_n s_k &= 0 \quad (k = 0, 1, 2 \ldots).
\end{aligned}$$

Eine explizite Darstellung der Potenzsummen s_ν geben die *Waringschen Formeln* (vgl. S. 219, vorletzte Zeile):

$$s_\nu = \nu \sum \frac{(\lambda_1 + \lambda_2 + \cdots + \lambda_n - 1)!}{\lambda_1!\,\lambda_2! \ldots \lambda_n!} \left(-\frac{a_1}{a_0}\right)^{\lambda_1} \cdot \left(-\frac{a_2}{a_0}\right)^{\lambda_2} \cdots \left(-\frac{a_n}{a_0}\right)^{\lambda_n};$$

das Summenzeichen ist über alle ganzen positiven Zahlen einschließlich 0 zu erstrecken, die der Gleichung $\lambda_1 + 2\lambda_2 + 3\lambda_3 + \cdots + n\lambda_n = \nu$ genügen.

Die Gleichungskoeffizienten drücken sich durch die Potenzsummen s in der Form:

$$a_\nu = a_0 \sum \frac{(-1)^{\lambda_1 + \lambda_2 + \cdots + \lambda_\nu}}{\lambda_1!\,\lambda_2! \ldots \lambda_\nu!} \left(\frac{s_1}{1}\right)^{\lambda_1} \cdot \left(\frac{s_2}{2}\right)^{\lambda_2} \cdots \left(\frac{s_\nu}{\nu}\right)^{\lambda_\nu}$$

aus; hierbei ist das Summenzeichen über alle ganzen positiven Zahlen einschließlich 0 zu erstrecken, die der Gleichung

$$\lambda_1 + 2\lambda_2 + 3\lambda_3 + \cdots + \nu\lambda_\nu = \nu$$

genügen.

In Determinantenform erhält man:

$$(-1)^\nu \cdot (\nu!)\, a_\nu = a_0 \cdot \begin{vmatrix} s_1 & 1 & 0 & \ldots 0 & 0 \\ s_2 & s_1 & 2 & \ldots 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ s_{\nu-1} & s_{\nu-2} & s_{\nu-3} & \cdots s_1 & \nu-1 \\ s_\nu & s_{\nu-1} & s_{\nu-2} & \cdots s_2 & s_1 \end{vmatrix}$$

In Verallgemeinerung der Potenzsummen s_ν kann man die symmetrische Funktion

$$\sum \varphi(\alpha_1) = \varphi(\alpha_1) + \varphi(\alpha_2) + \cdots + \varphi(\alpha_n)$$

betrachten, wobei $\varphi(z)$ irgendeine ganze rationale Funktion von z bedeutet und $\alpha_1, \alpha_2, \ldots, \alpha_n$ die Wurzeln der vorgelegten Gleichung $f(z) = 0$ des n^{ten} Grades sind. $\sum\varphi(\alpha_1)$ ist der Koeffizient von z^{-1} bei der Entwicklung $\frac{f'(z)}{f(z)}\varphi(z)$ nach fallenden Potenzen von z, wobei $f'(z)$ die erste Abgeleitete von $f(z)$ ist. (Cauchy, *Exercices de math.* **1** (1826), *Œuvres* (2) **6**, 401.) Wie $\frac{f'(z)}{f(z)}\varphi(z)$ eine erzeugende Funktion für die symmetrische Funktion $\sum\varphi(\alpha_1)$ ist, haben Borchardt, *Monatsb. d. Berl. Akad.* (1855), 165, *Ges. Werke*, Berlin 1888, S. 99 und Kronecker, *Monatsb. d. Berl. Akad.* (1880), 936 die typischen symmetrischen Funktionen der Gleichungswurzeln als Entwicklungskoeffizienten einer erzeugenden Funktion dargestellt.

Die symmetrischen Funktionen genügen gewissen partiellen Differentialgleichungen; vgl. Faà di Bruno, *Binäre Formen*, Leipzig 1881, S. 18, Netto, *Algebra* **1**, 133.

Tabellen, welche alle symmetrischen Funktionen der Wurzeln bis zum zehnten Grade für eine beliebige algebraische Gleichung enthalten, findet man bereits in Meyer Hirschs *Sammlung von Beispielen, Formeln u. Aufgaben aus der Buchstabenrechnung u. Algebra* (1804), des weiteren bei Cayley, *Coll. math. papers* **2**, 417 (ebenda in den Notes, S. 602 zahlreiche Literatur), Faà di Bruno, *Binäre Formen*, S. 302, Kostka, *Math.-Ver.* **16**. 429 (1907).

§ 6. Resultante und Diskriminante.

Zwei Gleichungen:

$$f(z) \equiv a_0 z^m + a_1 z^{m-1} + a_2 z^{m-2} + \cdots + a_m = 0,$$
$$g(z) \equiv b_0 z^n + b_1 z^{n-1} + b_2 z^{n-2} + \cdots + b_n = 0$$

haben dann und nur dann wenigstens eine gemeinsame Wurzel, wenn eine gewisse ganze rationale Funktion der Koeffizienten, welche die *Resultante der zwei Gleichungen* heißt, verschwindet. Da die charakteristische Bedingung für das Vorhandensein einer gemeinsamen Wurzel von $f(z) = 0$ und $g(z) = 0$ die Existenz eines gemeinsamen Teilers von $f(z)$ und $g(z)$ ist, kann man auch sagen: *Das Verschwinden der Resultante ist die notwendige und hinreichende Bedingung dafür, daß $f(z)$ und $g(z)$ einen gemeinsamen Teiler besitzen.* Auf Grund der letzten Aussage kann man die Resultante auch ohne Voraussetzung der Wurzelexistenz definieren und beim Beweise des Fundamentalsatzes der Algebra verwerten. Vgl. den Gauß-Gordanschen Beweis des Fundamentalsatzes der Algebra (Gauß, *Zweiter Beweis des Fundamentalsatzes*, Gordan, *Math. Ann.* **10**, 572 (1876), *Vorl. über Invariantentheorie*, **1**, 166).

Die Resultante R_{fg} läßt sich aus einem System von $m+n$ linearen homogenen Gleichungen als Determinante $(m+n)^{\text{ten}}$ Grades finden, nämlich

$$R_{fg} = \begin{vmatrix} a_0 & a_1 & a_2 & \cdots & a_m & 0 & 0 & \cdots & 0 \\ 0 & a_0 & a_1 & \cdots & a_{m-1} & a_m & 0 & \cdots & 0 \\ 0 & 0 & a_0 & \cdots & a_{m-2} & a_{m-1} & a_m & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & & \cdots & \cdots & \cdots & & a_m \\ b_0 & b_1 & b_2 & \cdots & b_n & 0 & 0 & \cdots & 0 \\ 0 & b_0 & b_1 & \cdots & b_{n-1} & b_n & 0 & \cdots & 0 \\ \vdots & & & & & & & & \\ 0 & 0 & 0 & \cdots & \cdots & \cdots & \cdots & & b_n \end{vmatrix} \begin{matrix} \\ \\ \\ \\ \\ \\ \\ \\ \end{matrix} \begin{matrix} \\ \\ \\ \\ \} n \text{ Zeilen} \\ \\ \\ \\ \\ \} m \text{ Zeilen} \end{matrix}$$

Die fraglichen linearen Gleichungen ergeben sich

a) durch das Verfahren von Euler (*Introductio in analysin infinitorum*, Lausanne 1748, t. **2**, Cap. 19, p. 265), das übrigens schon auf Leibniz (*Ges. Werke*, herausg. von Pertz, 3. Folge: *Math.*, herausg. von Gerhardt, Halle 1863, Bd. 7, S. 6)

zurückgeht. Man stellt die Bedingung dafür auf, daß zwei ganze Funktionen $F(z)$ und $G(z)$ vom Grade $\leqq m-1$ bezw. $\leqq n-1$ existieren und $fG+gF$ identisch verschwindet (vgl. S. 259). Eine Transposition der sich ergebenden linearen homogenen Gleichungen führt zu dem linearen homogenen System mit der Determinante R_{fg};

b) durch die von Sylvester sogenannte *diälytische* Methode. Diese eliminiert $1, z, z^2, \ldots, z^{n+m-1}$ aus den Gleichungen $f(z)=0$, $zf(z)=0$, $z^2f(z)=0, \ldots, z^{n-1}f(z)=0$, $g(z)=0$, $zg(z)=0, \ldots, z^{m-1}g(z)=0$ (Sylvester, *Phil. Mag.* **16**, (1840), *Cambr. Math. Journ.* **2** (1841), *Coll. math. papers* **1**, 54, 61, vgl. die Würdigung Sylvesters durch Noether, *Math. Ann.* **50**, 136 (1898)).

Die Resultante R_{fg} ist eine ganze rationale Funktion n^{ten} Grades der Koeffizienten $a_0, a_1, \ldots, a_m$ und m^{ten} Grades von $b_0, b_1, b_2, \ldots, b_n$; sie ist als Funktion der sämtlichen Größen $a_0, a_1, \ldots, a_m, b_0, b_1, \ldots, b_n$ unzerlegbar und ferner isobar vom Gewicht mn. R_{fg} ist linear und homogen aus f und g komponierbar, also $R_{fg}=fP+gQ$, wobei P und Q ganze Funktionen sind.

Die Unzerlegbarkeit und die Fähigkeit der linearen homogenen Komposition aus f und g sind charakteristische Eigenschaften der Resultante (vgl. S. 273).

Für die Vertauschung von f und g gilt:

$$R_{fg}=(-1)^{mn}R_{gf}.$$

Die Resultante läßt sich auch als *symmetrische Funktion der Gleichungswurzeln* auffassen (sogenannte erste Eulersche Methode, Euler, *Histoire de l'acad. de Berlin*, année 1748, 234, Cramer, *Introduction à l'analyse des lignes courbes*, Genf 1750, 660). Es ist:

$$\begin{aligned}
R_{fg} &= a_0^n g(\alpha_1)g(\alpha_2)\ldots g(\alpha_m),\\
R_{fg} &= (-1)^{mn}\cdot b_0^m f(\beta_1)f(\beta_2)\ldots f(\beta_n),\\
R_{fg} &= a_0^n b_0^m(\alpha_1-\beta_1)(\alpha_1-\beta_2)\ldots(\alpha_1-\beta_n)\cdot\\
&\qquad (\alpha_2-\beta_1)(\alpha_2-\beta_2)\ldots(\alpha_2-\beta_n)\cdot\\
&\qquad\quad \vdots \qquad\quad \vdots \qquad\quad \vdots\\
&\qquad (\alpha_m-\beta_1)(\alpha_m-\beta_2)\ldots(\alpha_m-\beta_n);
\end{aligned}$$

hierbei sind $\alpha_1, \alpha_2, \ldots, \alpha_m$ die Wurzeln der Gleichung $f(z)=0$ und $\beta_1, \beta_2, \ldots, \beta_n$ diejenigen von $g(z)=0$.

Die zwei ganzen Funktionen $f(z)$ und $g(z)$ haben dann und nur dann einen größten gemeinsamen Teiler, der genau vom k^{ten} Grade ist, wenn R_{fg}, R_1, R_2, ..., R_{k-1} verschwinden, hingegen R_k von Null verschieden ist. R_j entsteht aus R_{fg}, indem man die j letzten Zeilen der a, die j letzten Zeilen der b und die $2j$ letzten Kolonnen streicht (vgl. Faà di Bruno, *Binäre Formen,* Leipzig 1881, S. 57, Scheibner, *Leipz. Ber.* (1888), 3, Noether, Lüroth in den *Sitzungsb. d. physik.-med. Societät in Erlangen* (1895), Lüroth, *Ztschr. f. Math. u. Phys.* **40**, 247 (1895), Heffter, *Math. Ann.* **54**, 541 (1901)).

Für $m = n$ kann R_{fg} in eine Determinante m^{ten} Grades zusammengezogen werden, nämlich:

$$R_{fg} = (-1)^{\frac{m(m+1)}{2}} \left| c_{ik} \right| \qquad (i,k = 0,1,2,\ldots,m-1).$$

Zu dieser sogen. *Bézoutschen Determinante* $\left| c_{ik} \right|$ gelangt man durch das abgekürzte Verfahren von Bézout (*Mém. Acad. de Paris* (1764), Jacobi, *Journ. f. Math.* **15**, 101 (1836), *Ges. Werke* **3**, 295, Cauchy, *Exercices d'analyse* (1840), abgedruckt *Nouvelles Annales* (2) **15**, 385 (1876)). Aus $f = a_0 z^m + a_1 z^{m-1} + \cdots + a_m$ und $g = b_0 z^m + b_1 z^{m-1} + \cdots + b_m$ bilde man sich

$$\begin{aligned}
b_0 f - a_0 g &= 0,\\
(b_0 z + b_1) f - (a_0 z + a_1) g &= 0,\\
(b_0 z^2 + b_1 z + b_2) f - (a_0 z^2 + a_1 z + a_2) g &= 0,\\
&\vdots\\
(b_0 z^{m-1} + b_1 z^{m-2} + \cdots + b_{m-1}) f - (a_0 z^{m-1} + a_1 z^{m-2} + \cdots + a_{m-1}) g &= 0.
\end{aligned}$$

Die erhaltenen Gleichungen:

$$c_{i0} z^{m-1} + c_{i1} z^{m-2} + \cdots + c_{im-1} = 0 \qquad (i = 0,1,2,\ldots,m-1)$$

sind sämtlich vom Grade $m - 1$; durch Elimination von 1, z, z^2, ..., z^{m-1} erhält man die Determinante $\left| c_{ik} \right|$ $(i,k = 0,1,2,\ldots,m-1)$. Sie ist *symmetrisch,* und es ist

$$\begin{aligned}
&c_{i0} = d_{i+1,0}, \quad c_{i1} = d_{i+2,0} + d_{i+1,1},\\
&c_{i2} = d_{i+3,0} + d_{i+2,1} + d_{i+1,2}, \ldots, c_{im-1} = d_{mi},
\end{aligned}$$

wobei $d_{ik} = a_i b_k - a_k b_i$ ist.

Die Größen c_{ik} ergeben sich nach Cayley (*Journ. f. Math.* **53**, 366 (1857), *Coll. math. papers* **4**, 38) als *Entwicklungs-*

koeffizienten der scheinbar gebrochenen Funktion $\frac{f(x)g(y)-g(x)f(y)}{y-x}$ nach ganzen positiven Potenzen von x und y. Es ist

$$\frac{f(x)g(y)-f(y)g(x)}{y-x}=\sum_{i=0}^{i=m-1}\sum_{k=0}^{k=m-1}c_{ik}x^{m-k-1}y^{m-i-1}$$

Der gefundenen ganzen rationalen Funktion kann man mit Sylvester (*Phil. Transact.* (1853), *Coll. math. papers* **1**, 430 u. 545) die quadratische Form $\sum_{i=0}^{i=m-1}\sum_{k=0}^{k=m-1}c_{ik}t_it_k$ der m Variablen $t_0, t_1, \ldots, t_{m-1}$ zuordnen; diese heißt die *Bézoutiante von f und g*.

Notwendig und hinreichend, damit f und g einen größten gemeinsamen Divisor genau vom k^{ten} Grade haben, ist das Verschwinden der Determinanten

$$\sum \pm c_{00}c_{11}\cdots c_{m-1\,m-1},\quad \sum \pm c_{00}c_{11}\cdots c_{m-2\,m-2},\cdots$$
$$\sum \pm c_{00}c_{11}\cdots c_{m-k\,m-k},$$

hingegen muß $\sum \pm c_{00}c_{11}\cdots c_{m-k-1,\,m-k-1}$ von Null verschieden sein. *Die Bézoutiante von f und g hat dann genau den Rang* $m-k$.

Die voraufgehenden Betrachtungen lassen sich auch auf $m>n$ ausdehnen. Ist

$$f(z)=a_0z^m+a_1z^{m-1}+\cdots+a_m,\ g(z)=b_0z^n+b_1z^{n-1}+\cdots+b_n,$$

so gehe man zunächst von $f(z)$ und

$$g(z)=B_0z^m+B_1z^{m-1}+\cdots+B_{m-n-1}z^{n+1}+b_0z^n+b_1z^{n-1}+\cdots+b_n$$

aus und setze dann $B_0=B_1=\cdots=B_{m-n-1}=0$. Man erhält $R_{fg}=(-1)^{\frac{m(m+1)}{2}}\cdot\frac{1}{a_0^{m-n}}|c_{ik}|$ $(i,k=0,1,\ldots,m-1)$.

Der Fall, daß f und g verschiedene Grade haben, läßt sich auch durch Betrachtung von $f(z)=0$ und $z^{m-n}g(z)=0$ auf den Fall gleichen Grades zurückführen. Die erhaltene Resultante ist durch a_m^{m-n} zu dividieren, um die wahre Resultante R_{fg} zu finden.

Will man im Fall $m>n$ bei der Resultantenbildung *keine fremden Faktoren* einführen, sondern die wahre Resultante R_{fg} erhalten, so hat man aus den zwei Funktionen

$$f(z)=a_0z^m+a_1z^{m-1}+\cdots+a_m$$

und

$$g(z)=b_0z^n+b_1z^{n-1}+\cdots+b_n,$$

die n Gleichungen zu bilden: $fg_{n-s} - gf_{m-s} = 0$ $(s = n, n-1, \ldots, 1)$,

wobei $f_t = a_0 z^t + a_1 z^{t-1} + \cdots + a_t$ $(t = m-n, m-n+1, \ldots, m-1)$,

$g_t = b_0 z^t + b_1 z^{t-1} + \cdots + b_t$ $(t = 0, 1, 2, \ldots, n-1)$

ist. Für $m = n$ gehen sie in die früheren über. Für $m > n$ füge man noch $g(z) = 0$, $zg(z) = 0$, $\ldots$, $z^{m-n-1}g(z) = 0$ hinzu. Man hat dann im ganzen m Gleichungen, von denen jede höchstens vom Grade $m-1$ ist und aus denen man $1, z, z^2, \ldots, z^{m-1}$ eliminieren kann. Diese Behandlung des Falles $m > n$ nach der Methode von Bézout, bei der die Determinante jedoch *nicht* wie im Fall $m = n$ symmetrisch ausfällt, geht auf Rosenhain, *Journ. f. Math.* **28**, 269 (1844) (Gleichungen 81 a. a. O.) zurück. Eine eingehende Darstellung bei H. Weber, *Algebra* **1**, 179.

Sowohl die Resultante als auch die im folgenden zu besprechende Diskriminante genügen gewissen *partiellen Differentialgleichungen*, die Brioschi (*Journ. f. Math.* **53**, 372 (1857)) aufgestellt hat. Vgl. die eingehende Behandlung von Noether in Faà di Bruno, *Binäre Formen*, Leipzig 1881, S. 275. Die Literatur über Resultanten findet man bei Pascal, *Determinanten*, Leipzig 1900, S. 206, sowie bei E. Netto, *Enzyklopädie d. math. Wiss.* **1**, 245 (französische Bearbeitung von Le Vavasseur, *Encycl. des sc. math.* **1**, vol. 2, p. 73 ff.). Vgl. auch das Kapitel über Invariantentheorie.

Die Untersuchung der Resultante zweier ganzer Funktionen $f(z)$ und $g(z)$ je einer Variablen ist für das folgende *allgemeinste Problem aus der Theorie der algebraischen Gleichungen* grundlegend:

Gegeben seien r ganze Funktionen $f_1(z_1, z_2, \ldots, z_p)$, $f_2(z_1, z_2, \ldots, z_p)$, $\ldots$, $f_r(z_1, z_2, \ldots, z_p)$ der p Variablen $z_1, z_2, \ldots, z_p$. Man soll entscheiden, ob es Werte $z_1, z_2, \ldots, z_p$ gibt, die gleichzeitig alle Gleichungen $f_1(z_1, z_2, \ldots, z_p) = 0$, $f_2(z_1, z_2, \ldots, z_p) = 0, \ldots, f_r(z_1, z_2, \ldots, z_p) = 0$ befriedigen, und man soll, wenn es solche Werte gibt, diese bestimmen. Die Erledigung dieses Problems erfordert außer rationalen Operationen bloß die Behandlung von Gleichungen mit nur einer Unbekannten; es ist also auf den Fall $p = 1$ reduzierbar.

Sind die r Gleichungen $f_1 = 0$, $f_2 = 0, \ldots, f_r = 0$ miteinander verträglich, so heißt ein Wertsystem, das allen Gleichungen gleichzeitig genügt, eine *Wurzel* oder eine *Lösung des Gleichungssystems*.

Satz von Kronecker (*Festschrift, Journ. f. Math.* **92**, 30 (1882), *Ges. Werke* **2**, 280): *Der Gesamtinhalt eines beliebigen Systems von Gleichungen* $f_1(z_1, z_2, \ldots, z_p) = 0$, $f_2(z_1, z_2, \ldots, z_p) = 0$, $\ldots$, $f_r(z_1, z_2, \ldots, z_p) = 0$ *kann vollständig durch ein System von höchstens* $p + 1$ *Gleichungen* $\varphi_1(z_1, z_2, \ldots, z_p) = 0$, $\varphi_2(z_1, z_2, \ldots, z_p) = 0, \ldots, \varphi_{p+1}(z_1, z_2, \ldots, z_p) = 0$ *ersetzt werden, so daß jede Lösung des f-Systems eine solche des φ-Systems ist und umgekehrt.* Für das Resultat, daß unter Umständen tatsächlich $p + 1$ Gleichungen erforderlich sind, vgl. das von Vahlen, *Journ. f. Math.* **108**, 346 (1891) gegebene Beispiel einer algebraischen Raumkurve unseres R_3, die nur als Durchschnitt von vier Flächen darstellbar ist.

Hat man $p + 1$ allgemeine Gleichungen

$$\varphi_1(z_1, z_2, \ldots, z_p) = 0, \ \varphi_2(z_1, z_2, \ldots, z_p) = 0, \ldots,$$
$$\varphi_{p+1}(z_1, z_2, \ldots, z_p) = 0$$

von den Graden $m_1, m_2, \ldots, m_{p+1}$ mit Koeffizienten a, so gibt es eine ganze rationale Funktion der Koeffizienten a, die *sog. Resultante* $R(a)$ *des Gleichungssystems, die bis auf einen numerischen Faktor durch folgende zwei Eigenschaften eindeutig bestimmt ist: erstens* $R(a)$ *ist eine ganze rationale unzerlegbare Form der* a, *zweitens* $R(a)$ *komponiert sich linear und homogen aus* $\varphi_1, \varphi_2, \ldots, \varphi_{p+1}$, also $R(a) = P_1\varphi_1 + P_2\varphi_2 + \cdots + P_{p+1}\varphi_{p+1}$, wobei $P_1, P_2, \ldots, P_{p+1}$ ganze Funktionen von $z_1, z_2, \ldots, z_p$ sind. Die Resultante $R(a)$ enthält das Glied $a_{11}^{M_1} a_{22}^{M_2} \ldots a_{p+1,\,p+1}^{Mp+1}$, wobei $M_i = \dfrac{M}{m_i}$ $(i = 1, 2, \ldots, p+1)$, $M = m_1 m_2 \ldots m_{p+1}$ und a_{ii} $(i = 1, 2, \ldots, p)$ der Koeffizient von $z_i^{m_i}$ in $\varphi_i(z_1, z_2, \ldots, z_p)$, $a_{p+1,\,p+1}$ das konstante Glied in $\varphi_{p+1}(z_1, z_2, \ldots, z_p)$ ist; legt man dem Glied $a_{11}^{M_1} a_{22}^{M_2} \ldots a_{p+1,\,p+1}^{Mp+1}$ den Koeffizienten $+1$ bei, so ist die Resultante $R(a)$ eindeutig bestimmt.

Die Resultante $R(a)$ *ist eine homogene Funktion* M^{ten} *Grades der Koeffizientenreihe* a *von* φ_i, also eine homoge Funktion $\left(\sum_{i=1}^{i=p+1} M_i\right)^{\text{ten}}$ Grades aller Koeffizienten a. Die Resultante $R(a)$ ist eine *isobare Funktion des Gewichtes* M, dabei ist als Gewicht eines jeden Koeffizienten a aus φ_i die Zahl $m_i - s$ zu nehmen, wobei s die Summe aller Exponenten von $z_1, z_2, \ldots, z_p$ bedeutet, mit denen das betreffende a in der Funktion φ_i multipliziert ist.

Aus dem Theorem über das Gewicht ergibt sich der *Satz von Bézout* (*Théorie des équations algébriques*, Paris 1779): *Sind* $p+1$ *allgemeine Gleichungen* $\varphi_1(z_1, z_2, \ldots, z_{p+1}) = 0$, $\varphi_2(z_1, z_2, \ldots, z_{p+1}) = 0, \ldots, \varphi_{p+1}(z_1, z_2, \ldots, z_{p+1}) = 0$ *mit* $p+1$ *Unbekannten von den Graden* m_1 *bezw.* $m_2, \ldots, m_{p+1}$ *gegeben, so haben sie* $m_1 \cdot m_2 \cdots m_{p+1}$ *Lösungen.*

Als Literatur über die hier nur kurz angeschnittenen Fragen vgl.: Kronecker, *Festschrift, Journ. f. Math.* **92**, 1 (1882), *Ges. Werke* **2**, 237, Mertens, *Sitzungsb. der Wiener Akad.* **93**, Abt. 2, 527 (1886), **108**, Abt. 2a, 1173 u. 1344 (1899), Macaulay, *Proc. Lond. M. S.* **35**, 3 (1903), Netto, *Algebra* **2**, 1—203, sowie besonders Julius König, *Einleitung in die allgemeine Theorie der algebraischen Größen*, Leipzig 1903.

Bildet man von einer ganzen Funktion

$$f(z) = a_0 z^m + a_1 z^{m-1} + \cdots + a_m$$

und ihrer ersten Abgeleiteten

$$f'(z) = m a_0 z^{m-1} + (m-1) a_1 z^{m-1} + \cdots + a_{m-1}$$

die Resultante

$$R_{ff'} = \begin{vmatrix} a_0 & a_1 & a_2 & \ldots & a_m & 0 & \ldots & 0 \\ 0 & a_0 & a_1 & \ldots & a_{m-1} & a_m & \ldots & 0 \\ \vdots & & & & & & & \\ 0 & 0 & 0 & \ldots & \ldots & \ldots & \ldots & a_m \\ m a_0 & (m-1)a_1 & (m-2)a_2 & \ldots & a_{m-1} & 0 & \ldots & 0 \\ 0 & m a_0 & (m-1)a_1 & \ldots & 2a_{m-2} & a_{m-1} & \ldots & 0 \\ \vdots & & & & & & & \\ 0 & 0 & 0 & \ldots & \ldots & \ldots & \ldots & a_{m-1} \end{vmatrix}$$

($m-1$ Zeilen, m Zeilen)

so ist diese offenbar durch a_0 teilbar. $D = \frac{1}{a_0}(-1)^{\frac{m(m-1)}{2}} R_{ff'}$ heißt nach Sylvester (*Phil. Mag.* **2** (1851), 406, *Coll. math. papers* **1**, 280) die *Diskriminante* der algebraischen Gleichung $f(z) = 0$. Gauß (*Ges. Werke* **3**, 38) wendet hierfür die Bezeichnung „Determinante von $f(z)$" an.

Hat $f(z) = 0$ die m Wurzeln $\alpha_1, \alpha_2, \ldots, \alpha_m$, so ist:

$$D = (-1)^{\frac{m(m-1)}{2}} a_0^{m-2} f'(\alpha_1) f'(\alpha_2) \dots f'(\alpha_m),$$

$$D = a_0^{2m-2} \begin{vmatrix} 1 & 1 & \dots & 1 \\ \alpha_1 & \alpha_2 & \dots & \alpha_m \\ \alpha_1^2 & \alpha_2^2 & \dots & \alpha_m^2 \\ \vdots & \vdots & & \vdots \\ \alpha_1^{m-1} & \alpha_2^{m-1} & \dots & \alpha_m^{m-1} \end{vmatrix}^2$$

$$= a_0^{2m-2} \Delta(\alpha_1, \alpha_2, \dots, \alpha_m)^2,$$

wobei $\Delta(\alpha_1, \alpha_2, \dots, \alpha_m)$ das Differenzenprodukt der Gleichungswurzeln bedeutet (vgl. S. 68 unten und 69 oben).

Setzt man $s_\nu = \alpha_1^\nu + \alpha_2^\nu + \dots + \alpha_m^\nu$, ist also s_ν die Summe der ν^{ten} Potenzen der Gleichungswurzeln, so wird D durch eine rekurrierende Determinante (vgl. S. 67) gegeben:

$$D = a_0^{2m-2} \begin{vmatrix} s_0 & s_1 & \dots & s_{m-1} \\ s_1 & s_2 & \dots & s_m \\ s_2 & s_3 & \dots & s_{m+1} \\ \vdots & \vdots & & \vdots \\ s_{m-1} & s_m & \dots & s_{2m-2} \end{vmatrix}.$$

Die Diskriminante D ist eine ganze homogene Funktion der Gleichungskoeffizienten $a_0, a_1, \dots, a_m$ vom Grade $2m-2$; sie ist eine unzerlegbare Funktion der Koeffizienten, und ferner ist sie in ihnen isobar vom Gewicht $m(m-1)$. Das Verschwinden der Diskriminante ist die notwendige und hinreichende Bedingung dafür, daß die Gleichung $f(z) = 0$ wenigstens zwei gleiche Wurzeln besitzt.

Sei D_ν die rekurrierende Determinante:

$$\begin{vmatrix} s_0 & s_1 & \dots & s_{\nu-1} \\ s_1 & s_2 & \dots & s_\nu \\ \vdots & \vdots & & \vdots \\ s_{\nu-1} & s_\nu & \dots & s_{2\nu-2} \end{vmatrix},$$

so gilt folgender Satz: *Die Gleichung $f(z) = 0$ hat dann und*

nur dann genau k untereinander verschiedene Wurzeln, wenn die Determinanten D_{k+1}, D_{k+2}, ..., $D_m = \frac{D}{a_0^{2m-2}}$ *verschwinden, hingegen* D_k *von Null verschieden ist.* (Borchardt, *Journ. de math.* **12** (1847), *Ges. Werke*, Berlin 1888, S. 24, L. Baur, *Math. Ann.* **50**, 241 (1898).)

Bei einer Gleichung mit reellen Koeffizienten ohne mehrfache Wurzeln besitzt die Diskriminante positives oder negatives Vorzeichen, je nachdem die Gleichung eine gerade oder ungerade Anzahl von Paaren konjugiert imaginärer Wurzeln hat. (A. Brill, *Math. Ann.* **12**, 87 (1877).)

Deutet man bei einer Gleichung mit reellen Koeffizienten $a_0, a_1, \ldots, a_m$ diese als homogene Koordinaten im R_m, so heißt $D(a_0, a_1, \ldots, a_m) = 0$ die *Diskriminantenfläche.* Der Name stammt von Kronecker, *Monatsb. d. Berl. Akad.* (1878), 120, *Ges. Werke* **2**, 68. Vgl. ferner Hilbert, *Math. Ann.* **30**, 437 (1887).

Ebenso wie die Resultante ist auch die Diskriminante auf Gleichungssysteme erweitert worden. Vgl. die bei der Resultante von Gleichungssystemen zitierte Literatur auf S. 274.

§ 7. Transformation einer Gleichung. Tschirnhausentransformation. Mit der Tschirnhausentransformation zusammenhängende Normalformen. Funktionen mehrerer Wurzeln. Gleichung der quadrierten Wurzeldifferenzen.

Ersetzt man in der Gleichung

$$f(z) = a_0 z^n + a_1 z^{n-1} + a_2 z^{n-2} + \cdots + a_n = 0$$

mit den Wurzeln $\alpha_1, \alpha_2, \ldots, \alpha_n$ die Größe z durch $t + h$, so hat die neue Gleichung n^{ten} Grades in t:

$$\frac{t^n f^{(n)}(h)}{n!} + \frac{t^{n-1} f^{(n-1)}(h)}{(n-1)!} + \cdots + f(h) = 0$$

die n Wurzeln $\alpha_1 - h$, $\alpha_2 - h, \ldots, \alpha_n - h$. Hierbei bedeutet $f^{(i)}(h)$ die i^{te} Abgeleitete von $f(h) = a_0 h^n + a_1 h^{n-1} + \cdots + a_n$.

Wählt man für h einen derartigen Wert, daß

$$\frac{f^{(n-1)}(h)}{(n-1)!} = n a_0 h + a_1$$

verschwindet, also $h = \frac{-a_1}{n a_0}$, *so fehlt der Gleichung in t ihr zweites Glied.*

Die Transformation $t = z - h$ ist nur ein Spezialfall der sog. *Tschirnhausentransformation* (vgl. S. 252). Unter dieser versteht man folgendes: t sei eine ganze rationale Funktion von z:

$$t = \varphi(z) = p_0 + p_1 z + p_2 z^2 + \cdots + p_{n-1} z^{n-1}.$$

Man soll die Gleichung in t aufstellen, welche die Wurzeln $t_1 = \varphi(\alpha_1)$, $t_2 = \varphi(\alpha_2)$, ..., $t_n = \varphi(\alpha_n)$ *besitzt, wenn* $f(z) = a_0 z^n + a_1 z^{n-1} + \cdots + a_n = 0$ *die Wurzeln* $\alpha_1, \alpha_2, \ldots, \alpha_n$ *hat.*[1]) Die Größen $t_i = \varphi(\alpha_i)$ $(i = 1, 2, \ldots, n)$ genügen der Gleichung

$$\Phi(t) = (t - t_1)(t - t_2) \ldots (t - t_n) =$$
$$= t^n + q_1 t^{n-1} + q_2 t^{n-2} + \cdots + q_n = 0.$$

Die Größen q_i $(i = 1, 2, \ldots, n)$ sind ganze homogene Funktionen i^{ten} Grades von $p_0, p_1, \ldots, p_{n-1}$, deren Koeffizienten ganze Funktionen von $\frac{a_1}{a_0}, \frac{a_2}{a_0}, \ldots, \frac{a_n}{a_0}$ sind. Die Gleichung $\Phi(t) = 0$ heißt eine *Tschirnhausenresolvente von* $f(z) = 0$. Man kann $q_1, q_2, \ldots, q_n$ auf folgende Weise finden: Man bilde sich durch sukzessive Potenzierung die Gleichungskette:

$$t = p_0 + p_1 z + p_2 z^2 + \cdots + p_{n-1} z^{n-1},$$
$$t^2 = p_0' + p_1' z + p_2' z^2 + \cdots + p_{n-1}' z^{n-1},$$
$$t^3 = p_0'' + p_1'' z + p_2'' z^2 + \cdots + p_{n-1}'' z^{n-1},$$
$$\vdots$$
$$t^n = p_0^{(n-1)} + p_1^{(n-1)} z + p_2^{(n-1)} z^2 + \cdots + p_{n-1}^{(n-1)} z^{n-1},$$

wobei die höheren als n^{ten} Potenzen von z mit Hilfe der Gleichung $f(z) = 0$ beseitigt sind. Setzt man $\sigma_i = t_1^i + t_2^i + \cdots + t_n^i$ und $s_i = \alpha_1^i + \alpha_2^i + \cdots + \alpha_n^i$, so ist:

$$\sigma_i = n p_0^{(i-1)} + p_1^{(i-1)} s_1 + p_2^{(i-1)} s_2 + \cdots + p_{n-1}^{(i-1)} s_{n-1}$$
$$(i = 1, 2, \ldots, n).$$

Aus den Potenzsummen $\sigma_1, \sigma_2, \ldots, \sigma_n$ lassen sich die Koeffizienten $q_1, q_2, \ldots, q_n$ der Gleichung $\Phi(t) = 0$ bestimmen. (Vgl. S. 266.) Die Gleichung $\Phi(t) = 0$ ist das Eliminationsresultat von z aus $f(z) = 0$ und $t - \varphi(z) = 0$. Sind $t_1, t_2, \ldots, t_n$

---

1) Die Aufstellung einer Gleichung für $\varphi(\alpha_1)$, wobei $\varphi(\alpha_1)$ voraussetzungsgemäß eine ganze rationale Funktion von α_1, die höchstens vom Grade $n - 1$ ist, bedeutet, involviert keine Beschränkung; denn man kann jede ganze oder gebrochene rationale Funktion einer Wurzel α_1 der Gleichung $f(z) = 0$ in eine ganze rationale Funktion von α_1 verwandeln, die höchstens vom Grade $n - 1$ ist.

untereinander verschieden, so ergeben sich die zugehörigen $\alpha_1, \alpha_2, \ldots, \alpha_n$ aus linearen Gleichungen. Ist $t_1 = \varphi(\alpha_1) = \varphi(\alpha_2) = \cdots = \varphi(\alpha_e)$, also t_1 eine e-fache Wurzel von $\Phi(t) = 0$, so haben die zwei Gleichungen $f(z) = 0$ und $t_1 - \varphi(z) = 0$ die e Wurzeln $\alpha_1, \alpha_2, \ldots \alpha_e$ gemeinsam, die man demnach aus einer Gleichung e^{ten} Grades finden kann. Hat $f(z) = 0$ lauter verschiedene Wurzeln, so muß $e < n$ sein. Hieraus folgt: *Kann man eine Tschirnhausenresolvente $\Phi(t) = 0$ von $f(z) = 0$ auflösen, so ist $f(z) = 0$ hierdurch entweder vollständig gelöst oder seine Behandlung auf Gleichungen niedrigeren Grades zurückgeführt.*

Durch geeignete Wahl der Größen $p_0, p_1, \ldots, p_{n-1}$ kann man aus $f(z) = 0$ eine *Tschirnhausenresolvente $\Phi(t) = 0$ von einfacherer Struktur als $f(z) = 0$* ableiten.

Setzt man bei der *Gleichung dritten Grades*

$$a_0 z^3 + a_1 z^2 + a_2 z + a_3 = 0$$

$t = p_0 + p_1 z + p_2 z^2$ und wählt das Verhältnis $p_0 : p_1 : p_2$ aus den zwei Gleichungen $q_1 = 0$, $q_2 = 0$, von denen die erste linear, die zweite quadratisch ist, so erhält man als Tschirnhausenresolvente die reine Gleichung $t^3 + q_3 = 0$.

Die *Gleichung vierten Grades*

$$a_0 z^4 + a_1 z^3 + a_2 z^2 + a_3 z + a_4 = 0$$

wird durch die Tschirnhausentransformation: $t = p_0 + p_1 z + p_2 z^2 + p_3 z^3$ gelöst, indem man p_0, p_1, p_2, p_3 derartig bestimmt, daß die zwei Gleichungen $q_1 = 0$ und $q_3 = 0$ vom ersten und dritten Grade bestehen. Die Tschirnhausenresolvente $t^4 + q_2 t^2 + q_4 = 0$ ist dann durch Quadratwurzeln lösbar.

Für die allgemeinen Gleichungen höheren als vierten Grades kann die Tschirnhausentransformation keine Auflösungsmethoden, sondern nur *Normalformen* liefern. Bestimmt man in $t = p_0 + p_1 z + p_2 z^2 + \cdots + p_{n-1} z^{n-1}$ die Größen p so, daß die Gleichungen $q_1 = 0$ und $q_2 = 0$ vom ersten und zweiten Grade bestehen, so erfordert dies außer rationalen Operationen nur das Ziehen einer Quadratwurzel. Die alsdann aus

$$a_0 z^n + a_1 z^{n-1} + \cdots + a_n = 0$$

hervorgehende Gleichung

$$t^n + q_3 t^{n-3} + q_4 t^{n-4} + \cdots + q_n = 0$$

heißt eine *Hauptgleichung*. Als *Hauptgleichung* bezeichnet man

jede Gleichung n^{ten} Grades, in der keine Glieder mit der $(n-1)^{\text{ten}}$ und $(n-2)^{\text{ten}}$ Potenz der Unbekannten auftreten. (Klein, *Ikosaeder*, S. 142.)

Wählt man die Größen p derartig, daß gleichzeitig die drei Gleichungen ersten, zweiten und dritten Grades $q_1 = 0$, $q_2 = 0$, $q_3 = 0$ bestehen, so führt dies auf eine Gleichung sechsten Grades. Durch geeignete Verfügung über die p kann diese Gleichung sechsten Grades mit Hilfe von Quadratwurzeln auf eine kubische Gleichung zurückgeführt werden. Im besonderen läßt sich jede *Gleichung fünften Grades durch Ziehen von Quadratwurzeln und Auflösung einer kubischen Gleichung in die sog. Bring-Jerrardsche Normalform* $t^5 + q_4 t + q_5 = 0$ bringen. Setzt man $t = u\sqrt[4]{q_4}$, so erhält man eine Normalform $u^5 - u - A = 0$ mit nur einem Parameter. (Vgl. S. 254, sowie die Darstellung bei Weber, *Algebra* **1**, 260, ferner Rahts, *Math. Ann.* **28**, 34 (1887).)

Hat man eine Gleichung n^{ten} Grades

$$a_0 z^n + a_1 z^{n-1} + \cdots + a_n = 0$$

und wählt:

$$\begin{aligned}
p_0 &= a_0 u_1 + a_1 u_2 + a_2 u_3 + \cdots + a_{n-1} u_n, \\
p_1 &= \quad a_0 u_2 + a_1 u_3 + \cdots + a_{n-2} u_n, \\
p_2 &= \quad\quad a_0 u_3 + \cdots + a_{n-3} u_n, \\
&\vdots \qquad\qquad\qquad \vdots \\
p_{n-1} &= \quad\quad\quad\quad a_0 u_n,
\end{aligned}$$

wobei $u_1, u_2, \ldots, u_n$ Unbestimmte sind, so geht

$$t = p_0 + p_1 z + p_2 z^2 + \cdots + p_{n-1} z^{n-1}$$

über in

$$\begin{aligned}
t = a_0 u_1 + (a_0 z + a_1) u_2 + (a_0 z^2 + a_1 z + a_2) u_3 + \cdots + \\
+ (a_0 z^{n-1} + a_1 z^{n-2} + \cdots + a_{n-1}) u_n
\end{aligned}$$

und stellt, da $a_0 \neq 0$ ist, die allgemeinste ganze Funktion $(n-1)^{\text{ten}}$ Grades von z dar. Der Koeffizient von u_i ist der Faktor von u^{n-i} in der Entwicklung von $\frac{f(u) - f(z)}{u - z}$ nach ganzen positiven Potenzen von u. Bildet man sich die ganze Funktion von u:

$$\frac{f(u) - f(z)}{u - z} - \frac{1}{n} f'(u) = u^{n-2} F_0(z) + u^{n-3} F_1(z) + \cdots + F_{n-2}(z),$$

wobei $f'(u)$ die erste Abgeleitete von $f(u)$ ist, und wählt $t = u_2 F_0(z) + u_3 F_1(z) + \cdots + u_n F_{n-2}(z)$, so sind die Koeffizienten der Tschirnhausenresolvente $\Phi(t) = 0$ simultane Invarianten der zwei Funktionen $f(z)$ und

$$G(z) = u_2 - (n-2)u_3 z + \binom{n-2}{2} u_4 z^2 - \cdots \pm u_n z^{n-2}.$$

(Hermite, *C. R.* **46**, 961 (1858), *Œuvres* **2**, 30, vgl. Weber, *Algebra* **1**, 240.)

Die sich nur auf eine einzige Gleichungswurzel beziehende Tschirnhausentransformation läßt sich auf *Funktionen mehrerer Gleichungswurzeln* ausdehnen: $\varphi(\alpha_1, \alpha_2, \ldots, \alpha_n)$ sei irgendeine ganze rationale Funktion der Wurzeln $\alpha_1, \alpha_2, \ldots, \alpha_n$ der Gleichung $f(z) = a_0 z^n + a_1 z^{n-1} + \cdots + a_n = 0$. *Ist* $\varphi(x_1, x_2, \ldots, x_n)$ *eine* ϱ*-wertige Funktion* (vgl. S. 217) *der Größen* $x_1, x_2, \ldots, x_n$ *und sind* $\varphi_1 = \varphi(x_1, x_2, \ldots, x_n)$, $\varphi_2(x_1, x_2, \ldots, x_n)$, $\ldots$, $\varphi_\varrho(x_1, x_2, \ldots, x_n)$ *die* ϱ *mit* φ *konjugierten Funktionen, so sind* $\varphi_1(\alpha_1, \alpha_2, \ldots, \alpha_n)$, $\varphi_2(\alpha_1, \alpha_2, \ldots, \alpha_n)$, $\ldots$, $\varphi_\varrho(\alpha_1, \alpha_2, \ldots, \alpha_n)$ *die* ϱ *Wurzeln der Gleichung:*

$$\Phi(t) = (t - \varphi_1)(t - \varphi_2) \ldots (t - \varphi_\varrho) = \\ = t^\varrho + q_1 t^{\varrho-1} + q_2 t^{\varrho-2} + \cdots + q_\varrho = 0.$$

Die Größen $q_1, q_2, \ldots, q_\varrho$ *sind als symmetrische Funktionen der Gleichungswurzeln von* $f(z) = 0$ *ganze rationale Funktionen von* $\frac{a_1}{a_0}, \frac{a_2}{a_0}, \ldots, \frac{a_n}{a_0}$. Die Gleichung $\Phi(t) = 0$ heißt eine *transformierte Gleichung,* auch *Resolvente von* $f(z) = 0$. *Eine auf die angegebene Weise abgeleitete transformierte Gleichung kann bei einer Gleichung* $f(z) = 0$ *von höherem als vierten Grade, außer wenn* φ_1 *symmetrisch oder zweiwertig ist, nicht niedriger als* n^{ten} *Grades sein.* (Für die Gleichung vierten Grades vgl. § 8, vgl. ferner den Schluß des § 10.)

Von den Resolventen $\Phi(t) = 0$ sei hier noch die Gleichung für die *Quadrate der Wurzeldifferenzen einer Gleichung* $f(z) = 0$ angeführt. Mit ihr haben sich Waring (*Miscellanea analytica* (1762), *Meditationes algebraicae,* 3. ed., p. 30 (1782)) und Lagrange (1798), *Œuvres* **8**, 24 u. 140 beschäftigt. Wählt man für φ_1 die $\frac{n(n-1)}{2}$wertige Funktion $\varphi_1 = (x_1 - x_2)^2$, so genügen die $\frac{n(n-1)}{2}$ Größen $(\alpha_i - \alpha_k)^2$ der Gleichung:

$$\Phi(t) = t^\varrho + q_1 t^{\varrho-1} + q_2 t^{\varrho-2} + \cdots + q_\varrho = 0.$$

$\left(\varrho = \frac{n(n-1)}{2}.\right)$ Ist σ_i die Summe der i^{ten} Potenzen der Wurzeln dieser Gleichung und $s_i = \alpha_1^i + \alpha_2^i + \cdots + \alpha_n^i$, so ist

$$\sigma_i = n s_{2i} - (2i) s_1 s_{2i-1} + \binom{2i}{2} s_2 s_{2i-2} + \cdots + (-1)^i \binom{2i}{i} \frac{s_i s_i}{2}$$

und aus σ_i $(i = 1, 2, \ldots, \varrho)$ kann man $q_1, q_2, \ldots, q_\varrho$ finden (vgl. S. 266). Die Größe q_ϱ ist gleich $(-1)^\varrho \Delta^2$, wobei Δ das Differenzenprodukt der n Wurzeln $\alpha_1, \alpha_2, \ldots, \alpha_n$ von $f(z) = 0$ ist (vgl. S. 68). Die Gleichung der quadrierten Wurzeldifferenzen hatte früher vornehmlich den Zweck, eine untere Grenze für den absoluten Betrag der Wurzeldifferenzen der Ausgangsgleichung zu finden; doch kann man dies nach Cauchy (*Analyse algébrique* (1821), Note III, *Œuvres* (2) **3**, 398), ohne die Gleichung selbst aufzustellen. Ist nämlich g eine obere Grenze für die Wurzeln von $f(z) = 0$, d. h. sind alle Wurzeln von $f(z) = 0$ absolut genommen kleiner als eine positive Größe g, so ist der absolute Betrag

$$|\alpha_i - \alpha_k| > \frac{|\Delta|}{(2g)^{\frac{n(n-1)}{2} - 1}},$$

wobei $|\Delta|$ der absolute Betrag des Differenzenproduktes der n Wurzeln von $f(z) = 0$ ist.

§ 8. Die kubische und die biquadratische Gleichung. Reziproke Gleichungen.

Die *kubische Gleichung:*

$$a_0 z^3 + a_1 z^2 + a_2 z + a_3 = 0 \tag{1}$$

transformiert sich durch $z = t - \frac{a_1}{3a_0}$ in

$$a_0 t^3 + b_2 t + b_3 = 0.$$

Setzt man

$$p = \frac{b_2}{a_0} = \frac{3a_2 a_0 - a_1^2}{3a_0^2},$$

$$q = \frac{b_3}{a_0} = \frac{-9a_0 a_1 a_2 + 2a_1^3 + 27a_0^2 a_3}{27a_0^3},$$

so wird die Auflösung jeder kubischen Gleichung (1) auf die von (2): $t^3 + pt + q = 0$ zurückgeführt. Ist ε eine beliebige der zwei primitiven dritten Einheitswurzeln $-\frac{1}{2} \pm \frac{i}{2}\sqrt{3}$, so lauten die drei Wurzeln der Gleichung (2):

$$t_1 = u + v,$$
$$t_2 = u\varepsilon + v\varepsilon^2,$$
$$t_3 = u\varepsilon^2 + v\varepsilon;$$

hierbei bedeutet u einen beliebigen der drei Werte von

$$\sqrt[3]{-\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}},$$

die Größe v ist gleich $\sqrt[3]{-\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}}$, jedoch darf für v nicht willkürlich einer der drei Werte der dritten Wurzel gewählt werden, sondern es ist derjenige zu nehmen, der durch $v = -\frac{p}{3u}$ bestimmt ist. Die Lösung

$$t_1 = \sqrt[3]{-\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}} + \sqrt[3]{-\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}}$$

heißt die *Cardanische Formel.*

Die erste Auflösung der kubischen Gleichung stammt von Scipione del Ferro. Nach Cardanos *Artis magnae sive de regulis algebraicis liber unus* (1545) fand er sie im Jahre 1515 und teilte sie dem Floridus mit. In einem wissenschaftlichen Wettstreit stellte letzterer dem Nicolo Tartaglia im Jahre 1535 dreißig Aufgaben, die sämtlich auf kubische Gleichungen hinausliefen und die diesen zu einer erneuten Auffindung der Lösung der kubischen Gleichung führten. Von Tartaglia erfuhr Cardano die Lösung und machte sie gegen den Willen des Tartaglia in seiner *Ars magna* öffentlich bekannt. Cardano lehrte auch die Beseitigung des zweiten Gliedes aus einer allgemeinen kubischen Gleichung (vgl. Cantor, *Geschichte der Math.* **2**, 484, Zeuthen, *Geschichte der Math. im 16. und 17. Jahrhundert, Abh. zur Geschichte der Math.* **17**, Leipzig 1903, S. 81, Eneström, *Bibl. math.* (3) **7**, 38 (1906)).

Ist $\frac{q^2}{4} + \frac{p^3}{27} = 0$, so hat die Gleichung die drei Wurzeln

$$t_1 = 2\sqrt[3]{-\frac{q}{2}}, \quad t_2 = t_3 = -\sqrt[3]{-\frac{q}{2}}.$$

Sind *p und q reelle Größen,* so hat die Gleichung $t^3 + pt + q = 0$, falls

a) $\frac{q^2}{4} + \frac{p^3}{27} > 0$ ist, eine reelle und zwei konjugiert imaginäre Wurzeln. Wählt man für u die reelle Wurzel und demnach auch $v = -\frac{p}{3u}$ reell, so ist t_1 die reelle Wurzel.

b) Ist $\frac{q^2}{4} + \frac{p^3}{27} = 0$, so sind sowohl die einfache Wurzel t_1 als auch die Doppelwurzel t_2 reell.

c) Ist $\frac{q^2}{4} + \frac{p^3}{27} < 0$, so sind alle drei Wurzeln reell. Die obigen Formeln haben dann den Übelstand, die reellen Wurzeln auf imaginäre Weise zu liefern. Man bezeichnet den Fall c) von alters her als *casus irreducibilis,* ohne daß die Bezeichnung etwas mit irreduzibler Gleichung (§ 9) zu tun hätte. Der Fall c) erwies sich insofern nicht reduzibel, als der Satz gilt: *Für die allgemeine Gleichung dritten Grades mit reellen Koeffizienten und drei reellen Wurzeln kann es keine Lösungsform geben, welche die Wurzeln nur mittels reeller Radikale liefert* (Mollame, *Nap. Rendiconti* (2) **4**, 167 (1890), Hölder, *Math. Ann.* **38**, 307 (1891), Kneser, ebenda **41**, 344 (1893)).

Der Fall c) läßt sich ohne Durchgang durch das Imaginäre mit Hilfe der *trigonometrischen Funktionen* erledigen. Ist $\frac{q^2}{4} + \frac{p^3}{27} < 0$, so muß $p = -P$ sein, wobei P eine positive Größe ist. Definiert man: $\cos\varphi = -\frac{q}{2 \cdot \frac{P}{3} \cdot \sqrt{\frac{P}{3}}}$, wobei φ durch passende Wahl des Vorzeichens von $\sqrt{\frac{P}{3}}$ im ersten Quadranten genommen werden kann, so sind:

$$t_1 = 2\sqrt{\frac{P}{3}} \cos\frac{\varphi}{3}, \quad t_2 = 2\sqrt{\frac{P}{3}} \cos\frac{\varphi + 2\pi}{3},$$

$$t_3 = 2\sqrt{\frac{P}{3}} \cos\frac{\varphi + 4\pi}{3}$$

die drei reellen Wurzeln der Gleichung $t^3 - Pt + q = 0$. Der Kern dieser von Vieta (*Supplementum geometriae* (1593), *De aequationum recognitione et emendatione* (1615)) stammenden Methode basiert auf der trigonometrischen Relation

$$\cos^3\left(\frac{\varphi}{3}\right) - \frac{3}{4}\cos\left(\frac{\varphi}{3}\right) - \frac{\cos\varphi}{4} = 0.$$

Besonders interessant ist *Lagranges* (*Réflexions sur la ré-*

solution algébrique des équations (1770), *Œuvres* **3**, 217) Lösung der kubischen Gleichung. Sie beruht auf gruppentheoretischer Grundlage. Sei:

$$z^3 + a_1 z^2 + a_2 z + a_3 = 0$$

die gegebene Gleichung mit den Wurzeln α_1, α_2, α_3, so ist $(\alpha_1 + \varepsilon\alpha_2 + \varepsilon^2\alpha_3)^3$ eine zweiwertige Funktion der Wurzeln, wenn ε eine beliebige der zwei dritten Einheitswurzeln $-\frac{1}{2} \pm \frac{i}{2}\sqrt{3}$ ist. Die angegebene zweiwertige Funktion genügt der *quadratischen Resolvente:*

$$T^2 - (9a_1a_2 - 2a_1^3 - 27a_3)T + (a_1^2 - 3a_2)^3 = 0$$

mit den zwei Wurzeln:

$$(\alpha_1 + \varepsilon\alpha_2 + \varepsilon^2\alpha_3)^3 = \frac{1}{2}(9a_1a_2 - 2a_1^3 - 27a_3) + \frac{3i}{2}\sqrt{3D},$$

$$(\alpha_1 + \varepsilon^2\alpha_2 + \varepsilon\alpha_3)^3 = \frac{1}{2}(9a_1a_2 - 2a_1^3 - 27a_3) - \frac{3i}{2}\sqrt{3D}.$$

Beachtet man, daß $\alpha_1 + \alpha_2 + \alpha_3 = -a_1$ ist, so findet man:

$$3\alpha_1 = -a_1 + \sqrt[3]{\frac{1}{2}(9a_1a_2 - 2a_1^3 - 27a_3) + \frac{3i}{2}\sqrt{3D}} + \sqrt[3]{\frac{1}{2}(9a_1a_2 - 2a_1^3 - 27a_3) - \frac{3i}{2}\sqrt{3D}},$$

$$3\alpha_2 = -a_1 + \varepsilon^2\sqrt[3]{\frac{1}{2}(9a_1a_2 - 2a_1^3 - 27a_3) + \frac{3i}{2}\sqrt{3D}} + \varepsilon\sqrt[3]{\frac{1}{2}(9a_1a_2 - 2a_1^3 - 27a_3) - \frac{3i}{2}\sqrt{3D}},$$

$$3\alpha_3 = -a_1 + \varepsilon\sqrt[3]{\frac{1}{2}(9a_1a_2 - 2a_1^3 - 27a_3) + \frac{3i}{2}\sqrt{3D}} + \varepsilon^2\sqrt[3]{\frac{1}{2}(9a_1a_2 - 2a_1^3 - 27a_3) - \frac{3i}{2}\sqrt{3D}}.$$

D hat den Wert $a_1^2a_2^2 - 27a_3^2 - 4a_1^3a_3 + 18a_1a_2a_3 - 4a_2^3$ und ist die *Diskriminante* von $z^3 + a_1z^2 + a_2z + a_3 = 0$ oder die durch -27 dividierte Diskriminante der quadratischen Resolvente. Von den bei der Lösung der kubischen Gleichung auftretenden zwei dritten Wurzeln ist nur eine willkürlich, weil ihr Produkt $(\alpha_1 + \varepsilon\alpha_2 + \varepsilon^2\alpha_3)(\alpha_1 + \varepsilon^2\alpha_2 + \varepsilon\alpha_3) = a_1^2 - 3a_2$ ist.

Vom Standpunkt der Galoisschen Theorie kann man sagen: Die Galoissche Gruppe der allgemeinen Gleichung dritten

Grades ist die symmetrische Gruppe dritten Grades, also eine Gruppe der Ordnung 6. Nach Adjunktion von $\sqrt{D}$, also der Quadratwurzel aus der Diskriminante, wird sie die alternierende Gruppe dritten Grades von der Ordnung 3, und die Gleichung wird in dem durch Adjunktion von $\sqrt{D}$ erweiterten Körper eine einfache Abelsche Gleichung. Mithin ist alsdann zur vollständigen Lösung einer kubischen Gleichung nach der allgemeinen Theorie der Abelschen Gleichungen außer der Kenntnis der dritten Einheitswurzeln nur noch das Ziehen einer einzigen dritten Wurzel erforderlich. Dies zeigt auch die oben für die Gleichung dritten Grades gegebene Lösung, wenn man die Relation

$$\frac{3i}{2}\sqrt{3D} = +\frac{3}{2}(\varepsilon - \varepsilon^2)\sqrt{D}$$

beachtet und bedenkt, daß nach Kenntnis der dritten Wurzel

$$\sqrt[3]{\frac{1}{2}(9a_1a_2 - 2a_1^3 - 27a_3) + \frac{3}{2}(\varepsilon - \varepsilon^2)\sqrt{D}}$$

die andere dritte Wurzel sich als Produkt des reziproken Wertes in $a_1^2 - 3a_2$ ergibt.

Für die Lösung der kubischen Gleichung mittels invariantentheoretischer Methoden vgl. Cayley, *Coll. math. papers* **2**, 542, Faà di Bruno, *Binäre Formen*, Leipzig 1881, S. 211, Gordan, *Invariantentheorie* **2**, 175. Für die Auflösung durch Tschirnhausentransformation vgl. § 7. Für andere Methoden vgl. Matthießen, *Antike u. moderne Algebra*, Leipzig 1878, S. 362, ferner J. Lüroth, *Vorles. über numerisches Rechnen*, Leipzig 1900, S. 166, Klein in *Dycks Katalog. math Modelle*, München 1892, S. 3.

Die Gleichung *vierten Grades*, auch *biquadratische Gleichung* genannt,

$$a_0z^4 + a_1z^3 + a_2z^2 + a_3z + a_4 = 0 \tag{1}$$

transformiert sich durch $z = t - \frac{a_1}{4a_0}$ in

$$a_0t^4 + b_2t^2 + b_3t + b_4 = 0.$$

Setzt man $p = \frac{b_2}{a_0}$, $q = \frac{b_3}{a_0}$, $r = \frac{b_4}{a_0}$, so wird die Auflösung jeder Gleichung vierten Grades (1) reduziert auf

$$t^4 + pt^2 + qt + r = 0. \tag{2}$$

Zur Lösung von (2) bedient man sich der *kubischen Resolvente*

$$u^3 + \frac{p}{2}u^2 + \frac{p^2 - 4r}{16}u - \frac{q^2}{64} = 0. \tag{3}$$

Die Gleichung (3) habe die Wurzeln u_1, u_2, u_3, dann lauten die Wurzeln der Gleichung (2):

$$\begin{aligned} t_1 &= \sqrt{u_1} + \sqrt{u_2} + \sqrt{u_3}, \\ t_2 &= \sqrt{u_1} - \sqrt{u_2} - \sqrt{u_3}, \\ t_3 &= -\sqrt{u_1} + \sqrt{u_2} - \sqrt{u_3}, \\ t_4 &= -\sqrt{u_1} - \sqrt{u_2} + \sqrt{u_3}. \end{aligned}$$

Unter $\sqrt{u_1}$ und $\sqrt{u_2}$ sind beliebige Werte der zweideutigen Quadratwurzeln zu verstehen, hingegen ist für $\sqrt{u_3}$ derjenige Wert der Quadratwurzel zu wählen, der durch

$$\sqrt{u_3} = -\frac{q}{8\sqrt{u_1}\sqrt{u_2}}$$

bestimmt ist. Die Diskriminante D der biquadratischen Gleichung (2) ist

$$D = 16p^4r - 4p^3q^2 - 128p^2r^2 + 144pq^2r + 256r^3 - 27q^4,$$

sie ist gleich der mit 16^3 multiplizierten Diskriminante der kubischen Resolvente (3). Das Verschwinden von D bedeutet, daß die biquadratische Gleichung und ihre kubische Resolvente mehrfache Wurzeln besitzen.

Die erste Auflösung der biquadratischen Gleichung stammt von Luigi Ferrari und ist zugleich mit der Lösung der Gleichung dritten Grades von Cardano, dem Lehrer des Ferrari, in der *Ars magna* (1545) publiziert worden.

Hat die Gleichung (2) *reelle Koeffizienten p, q, r,* so ist $u_1u_2u_3 = \frac{q^2}{64}$ nicht negativ. Falls $q \neq 0$ ist, sind demnach drei Fälle zu unterscheiden:

a) u_1, u_2, u_3 sind positiv, die biquadratische Gleichung hat vier reelle Wurzeln.

b) u_1 ist positiv, u_2 und u_3 sind negativ, t_1 und t_2 werden konjugiert imaginär, ebenso t_3 und t_4.

c) u_1 ist positiv, u_2 und u_3 sind konjugiert imaginär, mithin werden $\sqrt{u_2}$ und $\sqrt{u_3}$ konjugiert imaginär, also t_1 und t_2 reell, t_3 und t_4 konjugiert imaginär.

Ist $q = 0$, so hat die kubische Resolvente eine verschwindende Wurzel, daher die biquadratische Gleichung zwei entgegengesetzt gleiche Wurzelpaare.

Ist $D < 0$, so hat die kubische Resolvente zwei konjugiert imaginäre Wurzeln, und die biquadratische Gleichung hat zwei reelle und zwei konjugiert imaginäre Wurzeln.

Ist $D > 0$, so hat die kubische Resolvente drei reelle Wurzeln, und die biquadratische Gleichung hat vier reelle oder zwei Paare konjugiert imaginärer Wurzeln. Die charakteristische Bedingung dafür, daß die Gleichung $z^4 + pz^2 + qz + r = 0$ vier reelle Wurzeln besitzt, von denen auch zwei (aber nicht mehr) untereinander gleich sein können, ist $D \geqq 0$, $p < 0$, $q^2 - 4r > 0$. Vgl. hierzu Klein, *Elementarmathematik von höherem Standpunkte aus*, Teil I, ausg. von Hellinger, Leipzig 1908, S. 220.

Lagrange (*Réflexions sur la résolution algébrique des équations, Œuvres* **3**, 266, vgl. oben S. 218) verwendet zur Auflösung der biquadratischen Gleichung:

$$z^4 + a_1 z^3 + a_2 z^2 + a_3 z + a_4 = 0$$

mit den Wurzeln α_1, α_2, α_3, α_4 die dreiwertige Funktion $(\alpha_1 + \alpha_2 - (\alpha_3 + \alpha_4))^2$. Sie und ihre zwei konjugierten Werte $(\alpha_1 + \alpha_3 - (\alpha_2 + \alpha_4))^2$ und $(\alpha_1 + \alpha_4 - (\alpha_3 + \alpha_2))^2$ genügen der kubischen Resolvente:

$$(3') \quad U^3 - (3a_1^2 - 8a_2)U^2 + (3a_1^4 - 16a_1^2 a_2 + 16a_2^2 + 16a_1 a_3 - 64a_4)U - (a_1^3 - 4a_1 a_2 + 8a_3)^2 = 0.$$

Bezeichnet man die Wurzeln der Gleichung (3') mit U_1, U_2, U_3, so ist:

$$4\alpha_1 = -a_1 + \sqrt{U_1} + \sqrt{U_2} + \sqrt{U_3},$$
$$4\alpha_2 = -a_1 + \sqrt{U_1} - \sqrt{U_2} - \sqrt{U_3},$$
$$4\alpha_3 = -a_1 - \sqrt{U_1} + \sqrt{U_2} - \sqrt{U_3},$$
$$4\alpha_4 = -a_1 - \sqrt{U_1} - \sqrt{U_2} + \sqrt{U_3}.$$

Unter $\sqrt{U_1}$ und $\sqrt{U_2}$ sind beliebige Werte der zweideutigen Quadratwurzeln zu verstehen; hingegen ist für $\sqrt{U_3}$ derjenige Wert der Quadratwurzel zu wählen, der durch

$$\sqrt{U_3} = \frac{4a_1 a_2 - a_1^3 - 8a_3}{\sqrt{U_1}\sqrt{U_2}}$$

bestimmt ist; denn

$$(\alpha_1 + \alpha_2 - (\alpha_3 + \alpha_4)) \cdot (\alpha_1 + \alpha_3 - (\alpha_2 + \alpha_4)) \cdot (\alpha_1 + \alpha_4 - (\alpha_3 + \alpha_2))$$

ist eine symmetrische Funktion der Wurzeln, die gleich $4a_1a_2 - a_1^3 - 8a_3$ ist.

Für $a_1 = 0$, $a_2 = p$, $a_3 = q$, $a_4 = r$, $U = 16u$ geht die Gleichung (3′) in (3) über.

Dasselbe wie die obige dreiwertige Funktion leistet auch die ebenfalls dreiwertige Funktion $\alpha_1\alpha_2 + \alpha_3\alpha_4$ (Lagrange, *Œuvres* **3**, 263). Verwendet man an ihrer Stelle die ebenfalls dreiwertige Funktion $\frac{a_2}{3} - \alpha_1\alpha_2 - \alpha_3\alpha_4$, so gelangt man zu der kubischen Resolvente

$$(3'') \qquad T^3 - \frac{A}{3}T + \frac{B}{27} = 0$$

mit den Wurzeln $\frac{a_2}{3} - \alpha_1\alpha_2 - \alpha_3\alpha_4$, $\frac{a_2}{3} - \alpha_1\alpha_3 - \alpha_2\alpha_4$ und $\frac{a_2}{3} - \alpha_1\alpha_4 - \alpha_2\alpha_3$. Hierbei sind

$$A = a_2^2 + 12a_4 - 3a_1a_3,$$
$$B = 27a_1^2a_4 + 27a_3^2 + 2a_2^3 - 72a_2a_4 - 9a_1a_2a_3$$

die erste und zweite Invariante der biquadratischen Gleichung:

$$z^4 + a_1z^3 + a_2z^2 + a_3z + a_4 = 0;$$

ihre Diskriminante, die gleich derjenigen der Resolventen (3′) und (3″) ist, hat den Wert $D = \frac{1}{27}(4A^3 - B^2)$. Diese Methode beruht auf invariantentheoretischer Grundlage, vgl. Cayley, *Coll. math. papers* **2**, 545, Faà di Bruno, *Binäre Formen*, S. 214, Gordan, *Invariantentheorie* **2**, 191.

Vom Standpunkt der Galoisschen Theorie kann man sagen: Die Galoissche Gruppe der allgemeinen Gleichung vierten Grades ist die symmetrische Gruppe vierten Grades, also eine Gruppe der Ordnung 24. Nach Adjunktion von $\sqrt{D}$, also der Quadratwurzel aus der Diskriminante, wird sie die alternierende Gruppe vierten Grades von der Ordnung 12. Diese hat eine invariante Untergruppe $\mathfrak{G}_4$ des Index 3 (vgl. S. 209). Eine Funktion der Wurzeln, die bei dieser Gruppe ungeändert bleibt, ist $(\alpha_1 - \alpha_2)(\alpha_3 - \alpha_4)$. In dem durch $\sqrt{D}$ erweiterten Körper genügt diese Funktion mit ihren in bezug auf die alternierende Gruppe konjugierten Werten $(\alpha_1 - \alpha_3)(\alpha_4 - \alpha_2)$ und $(\alpha_1 - \alpha_4)(\alpha_2 - \alpha_3)$ der Resolvente:

$$T^3 - AT + \sqrt{D} = 0,$$

die eine einfache Abelsche Gleichung ist und die Diskriminante B^2 besitzt. Zur Lösung dieser kubischen Normalgleichung ist die Kenntnis der dritten Einheitswurzeln und das Ziehen einer einzigen dritten Wurzel erforderlich. Ist dies geschehen, so hat die vorgelegte biquadratische Gleichung in dem erweiterten Körper die $\mathfrak{G}_4$ zur Galoisschen Gruppe. Da die $\mathfrak{G}_4$ eine invariante Untergruppe der Ordnung 2 besitzt, so hat man, um die Wurzeln der biquadratischen Gleichung zu finden, schließlich noch zwei Quadratwurzeln zu ziehen. Diese Operationen kommen auch in den anderen Lösungen zum Ausdruck; denn die kubischen Resolventen (3), (3′) und (3″) sind nach Adjunktion von $\sqrt{D}$ einfache Abelsche Gleichungen, und außer ihrer Lösung ist stets noch das Ziehen zweier Quadratwurzeln erforderlich, um die Wurzeln der vorgelegten biquadratischen Gleichung zu erhalten. — Die Auflösung der Gleichung vierten Grades läßt sich auch mit der Irrationalität des Oktaeders oder nach Adjunktion der Quadratwurzel der Diskriminante mit der Irrationalität des Tetraeders in Verbindung bringen. $\mathfrak{O}_{24}$ und $\mathfrak{T}_{12}$ sind ja mit der symmetrischen bezw. alternierenden Gruppe von vier Symbolen isomorph. (Vgl. S. 238 u. 240, sowie Klein, *Ikosaeder*, S. 126 u. Gordan, *Invariantentheorie* **2**, 213.)

Wie sich die Gleichung dritten Grades mittelst trigonometrischer Funktionen, so läßt sich die Gleichung vierten Grades mittelst elliptischer Funktionen behandeln. Hermite, *C. R.* **46**, 715, 961 (1858), *Œuvres* **2**, 22 und 30; vgl. auch C. Jordan, *Traité des substitutions*, p. 370. Für andere Methoden der Auflösung der Gleichung vierten Grades vgl. Matthießen, *Antike u. moderne Algebra*, S. 540, Tortolini, *Annali di mat.* **1**, 310 (1858). Für die Auflösung durch Tschirnhausentransformation vgl. § 7.

Auf Gleichungen niedrigeren Grades lassen sich auch die von Moivre (Miscellanea analytica (1730)) zuerst untersuchten und von Euler (1732) sogenannten *reziproken Gleichungen* zurückführen. (Historisches bei Tropfke, *Geschichte der Elementarmathematik* **1**, 265). *Als reziproke Gleichungen bezeichnet man solche, welche die Eigenschaft haben, neben jeder Größe λ auch den reziproken Wert $\frac{1}{\lambda}$ zur Wurzel zu haben. Eine Gleichung* $a_0 z^n + a_1 z^{n-1} + \cdots + a_n = 0$ *ist dann und nur dann reziprok, wenn ihre Koeffizienten so beschaffen sind, daß* $a_i = a_{n-i}$ $(i = 0, 1, 2, \ldots, \frac{n}{2}$ bezw. $\frac{n-1}{2})$ *oder* $a_i = -a_{n-i}$ $(i = 0, 1, 2, \ldots, \frac{n}{2}$ bezw. $\frac{n-1}{2})$ *ist.*

Befreit man die vorgelegte reziproke Gleichung durch Division mit $z-1$ bezw. $z+1$ von den möglicherweise vorhandenen Wurzeln ± 1, so hat sie stets die Form:

$$(1)\quad b_0 z^{2\varrho} + b_1 z^{2\varrho-1} + \cdots + b_{\varrho-1} z^{\varrho+1} + b_\varrho z^\varrho + b_{\varrho-1} z^{\varrho-1} + b_{\varrho-2} z^{\varrho-2} + \cdots + b_0 = 0.$$

Dividiert man diese reziproke Gleichung geraden Grades durch z^ϱ und setzt $z + \frac{1}{z} = Z_1$, so erhält man eine Gleichung ϱ^{ten} Grades in Z_1, nämlich:

$$(2)\qquad c_0 Z_1^\varrho + c_1 Z_1^{\varrho-1} + \cdots + c_\varrho = 0.$$

Zur Umformung dient hierbei die Rekursionsformel

$$Z_{i+1} = Z_1 Z_i - Z_{i-1} \qquad (i = 1, 2, \ldots),$$

wobei $Z_i = z^i + \frac{1}{z^i}$ $(i = 1, 2, \ldots)$ und $Z_0 = 2$ ist. Aus jeder der ϱ Wurzeln von (2) gehen zwei Wurzeln von (1) hervor; sie ergeben sich aus der quadratischen Gleichung $z^2 - z Z_1 + 1 = 0$.

§ 9. Körper. Reduzibilität und Irreduzibilität. Algebraischer Körper. Primitive und imprimitive Größen. Normalgleichung.

Unter einem *Körper, Rationalitätsbereich* oder *Feld* wird im folgenden jedes *unendliche* System von Größen verstanden, das von der Vollständigkeit ist, daß die Summe, die Differenz, das Produkt und der Quotient irgend zweier Größen des Systems immer wieder zu Größen desselben Systems führt (vgl. S. 176, wo auch endliche Körper betrachtet wurden). Jeder auf diese Weise definierte Körper enthält alle rationalen Zahlen. *Der Körper aller rationalen Zahlen ist der kleinste mögliche Körper;* er heißt der *absolute Rationalitätsbereich.* Ein Körper kann mit den rationalen Zahlen erschöpft sein oder außer ihnen noch andere Größen $R, R', \ldots$ enthalten; in letzterem Fall umfaßt der Körper seiner Definition entsprechend auch alle Größen, die aus $R, R', \ldots$ durch die vier Grundoperationen hervorgehen. Erweitert man einen Körper durch Hinzufügung einer neuen, ihm bisher nicht angehörigen Größe, so spricht man von einer *Adjunktion.*

Der Begriff des Körpers findet sich bereits bei Abel, *Œuvres* **1**, 479; **2**, 220 und Galois, *Œuvres*, p. 34. Die Bezeichnung „Körper" ist von Dedekind eingeführt worden, vgl. Dirichlet-Dedekind, *Vorlesungen über Zahlentheorie,* 4. Aufl., Braunschweig 1894, S. 452. Von Kronecker (*Fest-*

schrift, Journ. f. Math. **92**, 3 (1882), *Ges. Werke* **2**, 248) stammt die Bezeichnung „Rationalitätsbereich". Hierunter versteht er jedoch in engerer Weise als oben nur solche Körper, die aus dem absoluten Rationalitätsbereich durch Adjunktion einer endlichen Anzahl von Konstanten und Unbestimmten hervorgehen (vgl. Koenig, *Algebraische Größen*, S. 183).

Hat man eine beliebige ganze Funktion von einer oder mehreren Variablen, so ist durch ihre Koeffizienten stets ein gewisser Körper P_1 mitgegeben, nämlich der kleinste Körper, der alle Koeffizienten der Funktion enthält.

Eine ganze Funktion mit Koeffizienten aus irgendeinem Körper P[1]) heißt *in diesem Körper reduzibel*, wenn sie sich in das Produkt zweier ganzer Funktionen zerlegen läßt, von denen jede von niedrigerem Grade als die ursprüngliche Funktion ist und ebenso wie diese nur Koeffizienten aus P besitzt. Kann eine ganze Funktion mit Koeffizienten aus P nicht derartig zerlegt werden, so heißt sie *in* P *irreduzibel*. Die Reduzibilität bezw. Irreduzibilität einer ganzen Funktion hängt wesentlich von P ab. Eine im *Körper aller Zahlen* irreduzible Funktion heißt *absolut irreduzibel* (vgl. S. 260). Eine ganze Funktion einer *einzigen* Variablen von höherem als erstem Grade ist im Körper aller Zahlen stets reduzibel. Hat eine ganze Funktion $f(z)$ einer *einzigen* Variablen von höherem als zweitem Grade *reelle Koeffizienten*, so ist $f(z)$ im Körper *aller reellen* Zahlen stets reduzibel; denn jede ganze Funktion $f(z)$ mit reellen Koeffizienten läßt sich in ein Produkt von Polynomen ersten und zweiten Grades mit reellen Koeffizienten zerlegen.

Eine im Körper P *reduzible ganze Funktion von beliebig vielen Veränderlichen kann nur auf eine Art als ein Produkt von im Körper* P *irreduziblen ganzen Funktionen dargestellt werden. Hierbei gelten im Körper* P *irreduzible Funktionen, deren Quotienten konstant sind, als nicht verschieden.* Für Methoden, welche die Zerlegung einer gegebenen ganzen Funktion in irreduzible Faktoren für einen beliebigen algebraischen Körper (über die Wortbedeutung vgl. unten) durch eine endliche Anzahl von Schritten ausführen, vgl. *Kronecker, Festschrift, Journ. f. Math.* **92**, 11 (1882), *Ges. Werke* **2**, 256, Runge, *Journ. f. Math.* **99**, 89 (1886), Molk, *Acta math.* **6**, 10 (1885), Mandl, *Journ. f. Math.* **113**, 252 (1894), Hancock, *Ann. de l'éc. norm.* (3) **17**, 89 (1900), Weber, *Algebra* **2**, 563, Koenig, *Algebraische Größen*, S. 127.

1) Der Körper P muß entweder mit dem Körper P_1 zusammenfallen oder ihn umfassen.

Eine Gleichung $f(z) = 0$ heißt in einem Körper P irreduzibel, wenn es die ganze Funktion $f(z)$ ist.

Satz von Abel (*Œuvres* **1**, 480, *Ostwalds Klassiker* Nr. 111, S. 5): *Hat eine in einem Körper* P *irreduzible Gleichung* $f(z) = 0$ *mit irgendeiner anderen Gleichung* $\varphi(z) = 0$, *die ebenso wie* $f(z)$ *auch nur Koeffizienten aus* P *besitzt, eine Wurzel gemeinsam, so genügen alle Wurzeln von* $f(z) = 0$ *auch der Gleichung* $\varphi(z) = 0$; *alsdann ist* $\varphi(z)$ *durch* $f(z)$ *teilbar.*

Hieraus folgt: *Eine irreduzible Gleichung kann mit keiner Gleichung niedrigeren Grades, die Koeffizienten aus dem gleichen Körper besitzt, eine Wurzel gemeinsam haben. Eine irreduzible Gleichung kann keine mehrfachen Wurzeln besitzen. Die linken Seiten zweier irreduziblen Gleichungen mit einer gemeinsamen Wurzel können sich nur um einen konstanten Faktor unterscheiden.*

Über die Irreduzibilität gibt aus der Natur der Gleichungskoeffizienten folgendes Schoenemann-Eisensteinsches Theorem, gewöhnlich Eisensteinscher Satz genannt, Aufschluß (Schoenemann, *Journ. f. Math.* **32**, 100 (1846), Eisenstein ebenda **39**, 166 (1850), Schoenemann, Reklamation ebenda **40**, 188 (1850)): *Ist in einer Gleichung mit ganzzahligen Koeffizienten der Koeffizient der höchsten Potenz gleich* 1 *(oder allgemeiner eine zu der Primzahl* p *relativ prime ganze Zahl) und sind alle übrigen Koeffizienten durch die nämliche Primzahl* p *teilbar, der letzte aber nicht durch* p^2, *so ist die Gleichung im Körper der rationalen Zahlen irreduzibel.*

Der Beweis des Schoenemann-Eisensteinschen Satzes läßt sich mit Hilfe des Satzes von Gauß (*Disquisitiones arithmeticae*, Art. 42) erbringen: *Eine jede im Körper der rationalen Zahlen reduzible ganze Funktion*

$$f(z) = z^n + a_1 z^{n-1} + a_2 z^{n-2} + \cdots + a_n$$

mit ganzzahligen Koeffizienten $a_1, a_2, \ldots, a_n$ *kann nie anders in zwei Faktoren*

$$f_1(z) = z^{n_1} + b_1 z^{n_1-1} + b_2 z^{n_1-2} + \cdots + b_{n_1} \quad (n_1 > 0)$$

$$f_2(z) = z^{n_2} + c_1 z^{n_2-1} + c_2 z^{n_2-2} + \cdots + c_{n_2} \quad (n_2 > 0)$$

mit rationalen Zahlen zerlegt werden, als daß sämtliche Größen $b_1, b_2, \ldots, b_{n_1}, c_1, c_2, \ldots, c_{n_2}$ *ganzzahlig sind.*

An den Schoenemann-Eisensteinschen Satz anknüpfende Untersuchungen, die aus der *Teilbarkeit der Gleichungskoeffizienten* auf die Reduzibilität bezw. Irreduzibilität der Glei-

chung schließen, stammen von Koenigsberger, *Journ. f. Math.* **115**, 53 (1895), Netto, *Math. Ann.* **48**, 81 (1897), M. Bauer, *Journ. f. Math.* **128**, 298 (1905), **132**, 21 (1907), **134**, 15 (1908), Dumas, *Journ. de math.* (6) **2**, 191 (1908), Perron, *Math. Ann.* **60**, 448 (1905); Irreduzibilitätskriterien auf Grund von *Größenbeziehungen zwischen den Gleichungskoeffizienten* bei Perron, *Journ. f. Math.* **132**, 288 (1907).

Irgendeine Größe ζ heißt *in bezug auf einen Körper* P *algebraisch* oder eine *aus* P *stammende algebraische Zahl,* wenn ζ einer algebraischen Gleichung mit Koeffizienten aus P genügt. *Zu jeder aus* P *stammenden algebraischen Zahl* ζ *gehört eine in* P *irreduzible Gleichung* $J(z) = 0$ *mit Koeffizienten aus* P; *ihre linke Seite ist bis auf eine beliebige,* P *angehörige Konstante eindeutig bestimmt.* $J(z) = 0$ *ist die Gleichung niedrigsten Grades mit Koeffizienten aus* P, *die* ζ *zur Wurzel hat.* Ist n der Grad von $J(z) = 0$, so heißt ζ eine *aus* P *stammende algebraische Größe* n^{ten} *Grades.* Jede rationale Funktion $\frac{g(\zeta)}{h(\zeta)}$ von ζ mit Koeffizienten aus P kann mit Hilfe der irreduziblen Gleichung $J(z) = 0$ vom Grade n in eine ganze rationale Funktion

$$c_0 + c_1\zeta + c_2\zeta^2 + \cdots + c_{n-1}\zeta^{n-1}$$

mit Koeffizienten aus P umgewandelt werden. Durchlaufen $c_0, c_1, c_2, \ldots, c_{n-1}$ alle Größen aus P, so stellt

$$c_0 + c_1\zeta + c_2\zeta^2 + \cdots + c_{n-1}\zeta^{n-1}$$

alle Größen des aus P durch Adjunktion von ζ hervorgehenden Körpers dar, und zwar eine jede nur einmal. Der aus P durch Adjunktion von ζ hervorgehende Körper heißt ein *algebraischer Körper über* P (Weber, *Algebra* **1**, 492), ein *algebraischer Oberkörper von* P oder ein *Relativkörper in bezug auf* P (Hilbert, *Math.-Ver.* **4**, *die Theorie der algebraischen Zahlkörper*, S. 203 (1897)) oder ein dem *Bereich* P *entstammender Gattungsbereich* (Kronecker, *Festschrift,* S. 7, jedoch benützt Kronecker diese Bezeichnung nur bei dem von ihm verwendeten engeren Begriff des Rationalitätsbereiches). P heißt der *Stammbereich.* Genügt ζ einer in P irreduziblen Gleichung n^{ten} Grades, so heißt *der Körper* (P, ζ) *vom Grade n.* Ein *aus dem absoluten Rationalitätsbereich* hervorgehender algebraischer Körper heißt kurzweg ein *algebraischer Körper.* Dieser läßt sich auch so definieren: Er geht aus dem absoluten Rationalitätsbereich durch Adjunktion einer Wurzel einer irreduziblen Gleichung mit ganzzahligen Koeffizienten hervor.

Hat man eine beliebige endliche Anzahl $\eta_1, \eta_2, \ldots, \eta_r$ aus einem Körper P stammender algebraischer Zahlen, so kann man stets eine und sogar unendlich viele aus P stammende algebraische Zahlen ζ finden, so daß ζ eine ganze rationale Funktion von $\eta_1, \eta_2, \ldots, \eta_r$ mit Koeffizienten aus P wird und $\eta_1, \eta_2, \ldots, \eta_r$ ganze rationale Funktionen von ζ mit Koeffizienten aus P sind. *Jeder durch Adjunktion einer endlichen Anzahl aus* P *stammender algebraischer Größen aus* P *hervorgehende Körper über* P *kann daher entstanden gedacht werden durch Adjunktion einer einzigen Wurzel einer in* P *irreduziblen algebraischen Gleichung mit Koeffizienten aus* P (Satz von Abel, *Œuvres* **1**, 547, vgl. auch Galois, *Œuvres*, p. 37). Man kann ζ einfach in der Form $\mu_1\eta_1 + \mu_2\eta_2 + \cdots + \mu_r\eta_r$ wählen, wobei für $\mu_1, \mu_2, \ldots, \mu_r$ gewöhnliche rationale Zahlen genommen werden können (vgl. G. Cantor, *Math. Ann.* **5**, 133 (1872), Weber, *Algebra* **1**, 500).

Hat die im Körper P irreduzible Gleichung $J(z) = 0$ vom Grade n die Wurzeln $\xi_1, \xi_2, \ldots, \xi_n$ und ist $\varphi(\xi_1)$ eine ganze rationale Funktion von ξ_1 mit Koeffizienten aus P, so genügt $t_1 = \varphi(\xi_1)$ einer Gleichung

$$\Phi(t) = (t - t_1)(t - t_2)\cdots(t - t_n) = 0$$

mit den Wurzeln $t_i = \varphi(\xi_i)$ $(i = 1, 2, \ldots, n)$.

Die aus der irreduziblen Gleichung $J(z) = 0$ *hervorgehende Tschirnhausenresolvente* $\Phi(t) = 0$ (vgl. S. 277) *hat Koeffizienten aus* P *und ist entweder eine in* P *irreduzible Gleichung oder ihre linke Seite ist die* e^{te} *Potenz einer in* P *irreduziblen ganzen Funktion* $\Phi_1(t)$ *vom Grade* n_1. Hieraus folgt: *Die Größen*

$$\varphi(\xi_1), \varphi(\xi_2), \ldots, \varphi(\xi_n)$$

sind entweder sämtlich untereinander verschieden oder sie zerfallen in n_1 *Systeme von je e gleichen:*

$$\varphi(\xi_1) = \varphi(\xi_2) \cdots = \varphi(\xi_e), \varphi(\xi_{e+1}) = \varphi(\xi_{e+2}) = \cdots = \varphi(\xi_{2e}), \ldots,$$
$$\varphi(\xi_{n-e+1}) = \varphi(\xi_{n-e+2}) = \cdots = \varphi(\xi_n).$$

Die Größen $t_1 = \varphi(\xi_1)$, $t_2 = \varphi(\xi_2), \ldots, t_n = \varphi(\xi_n)$ *heißen in bezug auf den Körper* P *konjugierte Größen. Sind sie sämtlich voneinander verschieden, dann und auch nur dann ist nicht bloß* $t_i = \varphi(\xi_i)$ *eine rationale Funktion von* ξ_i, *sondern auch umgekehrt* ξ_i *eine rationale Funktion von* t_i *mit Koeffizienten aus* P.

Adjungiert man einem Körper P die algebraische aus P stammende Größe ξ_1, so erhält man einen Körper (P, ξ_1) über P

Adjungiert man statt ζ_1 eine rationale Funktion $t_1 = \varphi(\zeta_1)$ von ζ_1 mit Koeffizienten aus P, so erhält man einen Körper (P, t_1), der nur Größen aus dem Körper (P, ζ_1) enthält. Ist t_1 von allen seinen konjugierten Größen verschieden, dann und nur dann sind die Körper (P, t_1) und (P, ζ_1) identisch. In diesem Fall heißt $t_1 = \varphi(\zeta_1)$ eine *primitive Größe des Körpers* (P, ζ_1).

Ein Körper, der sämtliche Größen des Körpers P umfaßt und dessen Größen ausnahmslos dem algebraischen Körper (P, ζ_1) über P angehören, ohne jedoch den Körper (P, ζ_1) zu erschöpfen, heißt ein *Divisor*, echter *Teiler* oder *Unterkörper von* (P, ζ_1). Jeder Divisor von (P, ζ_1) wird aus P erhalten, indem man dem Körper P eine gewisse ganze Funktion $t_1 = \varphi(\zeta_1)$ mit Koeffizienten aus P adjungiert. Charakteristisch dafür, daß (P, t_1) ein Divisor von (P, ζ_1) ist, erweist sich, daß t_1 eine *imprimitive Größe des Körpers* (P, ζ_1) ist, d. h. $t_1 = \varphi(\zeta_1)$ nicht von allen seinen konjugierten Größen verschieden ist. *Ein Körper* (P, ζ_1), *der nur primitive Größen enthält, also keine Divisoren besitzt, heißt ein primitiver Körper.*

Ist $J(z) = 0$ die irreduzible Gleichung mit Koeffizienten aus P, der ζ_1 genügt, so ist der Körper (P, ζ_1) dann und nur dann ein primitiver Körper, wenn alle Tschirnhausenresolventen von $J(z) = 0$ mit Koeffizienten aus P ausnahmslos in P irreduzibel sind. Alsdann heißt die irreduzible Gleichung $J(z) = 0$ selbst eine *in* P *irreduzible primitive Gleichung.* Jede in P irreduzible, nicht primitive Gleichung heißt *imprimitiv. Jede in* P *irreduzible Gleichung von Primzahlgrad ist notwendig primitiv.*

Ist $J(z) = 0$ *eine in* P *imprimitive Gleichung* n^{ten} *Grades mit der Wurzel* ζ_1 *und bedeutet* $t_1 = \varphi(\zeta_1)$ *irgendeine imprimitive Größe des Körpers* (P, ζ_1), *die einer in* P *irreduziblen Gleichung* n_1^{ten} *Grades genügt, so wird die in* P *irreduzible Gleichung* $J(z) = 0$ *im Körper* (P, t_1) *reduzibel, und* ζ_1 *genügt in ihm einer irreduziblen Gleichung* e^{ten} *Grades, wobei* $n = n_1 e$ *ist.*

Anders ausgedrückt: *Jede im Körper* P *imprimitive Gleichung* $J(z) = 0$ *läßt sich durch Elimination von* t *aus zwei Gleichungen:*

$$t^{n_1} + c_1 t^{n_1 - 1} + c_2 t^{n_1 - 2} + \cdots + c_{n_1} = 0$$

und

$$z^e + \alpha_1(t) z^{e-1} + \alpha_2(t) z^{e-2} + \cdots + \alpha_e(t) = 0$$

finden; die erste Gleichung ist im Körper P, *die zweite im Körper* (P, t) *irreduzibel* (vgl. **Jordan**, *Traité*, p. 260, **H. Weber**, *Algebra* **1**, 501ff. u. 524ff.).

Der obige Satz, daß der Grad n_1 jedes Unterkörpers (P, t_1) eines imprimitiven Körpers (P, ζ_1) vom Grad n — nur solche Körper können ja Unterkörper haben — ein Divisor von n ist, ergibt sich als Korollar aus dem Kronecker-Kneserschen Satz (Hölder, *Math. Ann.* **38**, 309 (1891), Kneser, *Math. Ann.* **30**, 195 (1887), *Journ. f. Math.* **106**, 51 (1890), Landsberg, ebenda, **132**, 8 (1907)): *Sind ξ und ζ die Wurzeln zweier im Körper P irreduzibler Gleichungen vom Grade n_1 bezw. n und sind die Grade der irreduziblen Gleichungen, denen ξ nach Adjunktion von ζ und ζ nach Adjunktion von ξ genügen, e_1 bezw. e, so ist $n_1 e = n e_1$. Ist also $e < n$, so muß $e_1 < n_1$ sein, d. h. die zwei Gleichungen wirken aufeinander reduzierend.* (In dem obigen speziellen Satz ist ξ die im Körper (P, ζ) liegende Größe $t_1 = \varphi(\zeta)$ und daher $e_1 = 1$).

Ist $J(z) = 0$ eine im Körper P irreduzible Gleichung n^{ten} Grades mit den Wurzeln $\zeta_1, \zeta_2, \ldots, \zeta_n$, so heißen die Körper $(P, \zeta_1), (P, \zeta_2), \ldots (P, \zeta_n)$ *in bezug auf* P *konjugierte Körper.* Ein Körper, der mit allen seinen konjugierten übereinstimmt, heißt ein Galoisscher oder Normalkörper. Eine im Körper P irreduzible Gleichung $J(z) = 0$ heißt eine *Normalgleichung* oder Galoissche *Gleichung*, wenn ihre sämtlichen Wurzeln rationale Funktionen einer einzigen mit Koeffizienten aus P sind. *Bei einer Normalgleichung sind alle Wurzeln nicht nur rationale Funktionen einer bestimmten Wurzel mit Koeffizienten aus* P, *sondern jeder beliebigen. Die Normalgleichungen und auch nur sie allein erzeugen Normalkörper.*

Hat eine beliebige reduzible oder irreduzible Gleichung $f(z) = 0$ mit Koeffizienten aus einem Körper P *lauter verschiedene Wurzeln $\alpha_1, \alpha_2, \ldots, \alpha_n$, so kann man stets eine Größe ξ_1 von folgender Beschaffenheit finden: ξ_1 ist eine ganze rationale Funktion von $\alpha_1, \alpha_2, \ldots, \alpha_n$ mit Koeffizienten aus* P, *und jede der Größen $\alpha_1, \alpha_2, \ldots, \alpha_n$ läßt sich als ganze rationale Funktion von ξ_1 mit Koeffizienten aus* P *ausdrücken.* (Vgl. oben S. 294.) *Die in* P *irreduzible Gleichung, der ξ_1 genügt, ist eine Normalgleichung.* Jede in P irreduzible Gleichung, die eine solche Funktion ξ_1 zur Wurzel hat, heißt eine Galoissche *Resolvente von $f(z) = 0$. Alle* Galoisschen *Resolventen von $f(z) = 0$ sind von gleichem Grade und jede eine Tschirnhausenresolvente einer jeden anderen.* Die Rückführung jeder beliebigen Gleichung $f(z) = 0$ mit Koeffizienten aus dem Körper P auf eine Galoissche Resolvente ist das Fundament der im § 10 behandelten Galoisschen Theorie. (Galois, *Œuvres*, p. 36.)

Hat man eine Gleichung $f(z) = 0$ ohne mehrfache Wurzeln, so kann man sich die Koeffizienten einer Funktion ζ_1 der eben definierten Beschaffenheit und eine Galoissche Resolvente, der ζ_1 genügt, auf folgende Art ohne Kenntnis der Wurzeln von $f(z) = 0$ verschaffen:

$$\tau_1 = \mu_1\alpha_1 + \mu_2\alpha_2 + \cdots + \mu_n\alpha_n$$

sei eine lineare homogene Funktion der Wurzeln $\alpha_1, \alpha_2, \ldots, \alpha_n$ mit zunächst unbestimmten Koeffizienten $\mu_1, \mu_2, \ldots, \mu_n$. Die Funktion τ_1 nehme bei den $\nu = n!$ Permutationen der Größen $\alpha_1, \alpha_2, \ldots, \alpha_n$ die ν Werte $\tau_1, \tau_2, \ldots, \tau_\nu$ an. Man bilde sich die transformierte Gleichung

$$g(\tau) = (\tau - \tau_1)(\tau - \tau_2)\cdots(\tau - \tau_\nu) = 0$$

von $f(z) = 0$. Die Koeffizienten von $g(\tau) = 0$ sind als symmetrische Funktionen der Wurzeln von $f(z) = 0$ von diesen unabhängig, die Diskriminante von $g(\tau) = 0$ wird eine ganze rationale Funktion von $\mu_1, \mu_2, \ldots, \mu_n$, die nicht identisch verschwindet, da die Wurzeln von $f(z) = 0$ untereinander verschieden sind. Alsdann wähle man, was auf unendlich viele Arten möglich ist, für

$$\mu_1, \mu_2, \ldots, \mu_n$$

solche rationale Zahlen, daß die Diskriminante einen von Null verschiedenen Wert besitzt, also $g(\tau) = 0$ keine mehrfachen Wurzeln hat, d. h. die Größen $\tau_1, \tau_2, \ldots, \tau_\nu$ sämtlich voneinander verschieden sind (vgl. Weber, *Algebra* **1**, 147, Bachmann, *Math. Ann.* **18**, 452 (1881)). Sind $\mu_1, \mu_2, \ldots, \mu_n$ derartig bestimmt, so zerlege man $g(\tau)$ in seine in P irreduziblen Faktoren. Ist $G(\tau)$ ein *beliebiger* irreduzibler Faktor von $g(\tau)$, so ist $G(\tau) = 0$ eine Galoissche Resolvente von $f(z) = 0$ und *jede* Wurzel von $G(\tau) = 0$ eine Funktion ζ_1 der verlangten Beschaffenheit.

Wir führen zum Schluß noch einen Satz von Hilbert (*Journ. f. Math.* **110**, 122 (1892)) über die Irreduzibilität einer ganzen Funktion mehrerer Veränderlicher an:

Ist eine ganze rationale $F(x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_m)$ von $m + n$ Variablen $x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_m$ in irgendeinem algebraischen Körper irreduzibel, so ist es stets auf unendlich viele Weisen möglich, in dieser Funktion für $y_1, y_2, \ldots, y_m$ ganze rationale Zahlen einzusetzen, so daß hierdurch die vorgelegte Funktion in eine Funktion der n Variablen $x_1, x_2, \ldots, x_n$ übergeht, die in eben diesem Körper irreduzibel ist.

§ 10. Die Galoissche Gruppe einer Gleichung. Natürliche und akzessorische Irrationalitäten. Rationale Resolventen. Reduktion der Gruppe und Rückführung einer Gleichung auf eine Kette von Hilfsgleichungen. Allgemeine Gleichungen. Affektlose Gleichungen.

P sei ein beliebiger Körper und $f(z) = 0$ irgendeine Gleichung n^{ten} Grades mit Koeffizienten aus P. Von der Gleichung $f(z) = 0$ wird nur, und zwar im folgenden ausnahmslos, vorausgesetzt, daß ihre sämtlichen Wurzeln $\alpha_1, \alpha_2, \ldots, \alpha_n$ voneinander verschieden sind. Wir betrachten die Gesamtheit aller Gleichungen mit Koeffizienten aus P, die $\alpha_1, \alpha_2, \ldots, \alpha_n$ ganz und rational enthalten und die richtige Relationen sind, falls für $\alpha_1, \alpha_2, \ldots, \alpha_n$ die Wurzeln der vorgelegten Gleichung $f(z) = 0$ gesetzt werden. *Die Gesamtheit aller Permutationen unter den Wurzeln $\alpha_1, \alpha_2, \ldots, \alpha_n$, die jede zwischen diesen bestehende richtige Gleichung mit Koeffizienten aus* P *wieder in eine richtige Gleichung überführen, bildet eine Permutationsgruppe* $\mathfrak{G}$ *n^{ten} Grades; diese heißt die Galoissche Gruppe der Gleichung $f(z) = 0$.* Die Galoissche Gruppe $\mathfrak{G}$ hängt von der Gleichung $f(z) = 0$ und dem der Betrachtung zugrunde liegenden Körper P ab.

Von irgendeiner rationalen Funktion $R(\alpha_1, \alpha_2, \ldots, \alpha_n)$ der Wurzeln $\alpha_1, \alpha_2, \ldots, \alpha_n$ mit Koeffizienten aus P sagt man: sie bleibt bei einer Permutation $\begin{pmatrix} \alpha_1, \alpha_2 \ldots \alpha_n \\ \alpha_{i_1}, \alpha_{i_2} \ldots \alpha_{i_n} \end{pmatrix}$ *numerisch ungeändert,* wenn $R(\alpha_1, \alpha_2, \ldots, \alpha_n)$ und $R(\alpha_{i_1}, \alpha_{i_2}, \ldots, \alpha_{i_n})$ für die Wurzeln $\alpha_1, \alpha_2, \ldots, \alpha_n$ von $f(z) = 0$ den gleichen Wert annehmen. Bei numerischer Unveränderlichkeit brauchen $R(\alpha_1, \alpha_2, \ldots, \alpha_n)$ und $R(\alpha_{i_1}, \alpha_{i_2}, \ldots, \alpha_{i_n})$ nicht formal übereinzustimmen.

Aus der Definition der Galoisschen Gruppe folgt:

Besitzt eine rationale Funktion von $\alpha_1, \alpha_2, \ldots, \alpha_n$ mit Koeffizienten aus P *eine dem Körper* P *angehörige Zahl als Wert, wenn man für $\alpha_1, \alpha_2, \ldots, \alpha_n$ die Wurzeln von $f(z) = 0$ setzt, so bleibt die betreffende Funktion bei allen Permutationen der Galoisschen Gruppe* $\mathfrak{G}$ *numerisch ungeändert.*

Dieser Satz ist umkehrbar: *Bleibt eine rationale Funktion der Gleichungswurzeln $\alpha_1, \alpha_2, \ldots, \alpha_n$ mit Koeffizienten aus* P *bei allen Permutationen der Galoisschen Gruppe* $\mathfrak{G}$ *numerisch ungeändert, so ist ihr Wert eine dem Körper* P *angehörige Zahl.*

Bleiben alle richtigen Gleichungen mit Koeffizienten aus P, *die $\alpha_1, \alpha_2, \ldots, \alpha_n$ ganz und rational enthalten, bei allen Permu-*

tationen einer Permutationsgruppe $\mathfrak{H}$ n^{ten} Grades unter den Größen $\alpha_1, \alpha_2, \ldots, \alpha_n$ richtig, und hat ferner jede rationale Funktion von $\alpha_1, \alpha_2, \ldots, \alpha_n$ mit Koeffizienten aus P, die bei allen Permutationen von $\mathfrak{H}$ numerisch ungeändert bleibt, einen Wert aus P, so enthält $\mathfrak{H}$ alle Permutationen der Galoisschen Gruppe $\mathfrak{G}$ der Gleichung $f(z) = 0$, und $\mathfrak{H}$ ist mit $\mathfrak{G}$ identisch.

$\xi_1 = V(\alpha_1, \alpha_2, \ldots, \alpha_n)$ sei eine beliebige rationale Funktion von $\alpha_1, \alpha_2, \ldots, \alpha_n$ mit Koeffizienten aus P von der Beschaffenheit, daß sich die Größen $\alpha_1, \alpha_2, \ldots, \alpha_n$ als ganze rationale Funktionen von ξ_1 mit Koeffizienten aus P in der Form

$$\alpha_k = \chi_k(\xi_1) \qquad (k = 1, 2, \ldots, n)$$

ausdrücken lassen. Die in P irreduzible Gleichung, die ξ_1 zur Wurzel hat, sei die Gleichung $G(\xi) = 0$ vom Grade g; $G(\xi) = 0$ ist also eine Galoissche Resolvente (vgl. S. 297) von $f(z) = 0$. Bedeutet ξ_i eine beliebige der g Wurzeln von $G(\xi) = 0$, so sind $\chi_1(\xi_i), \chi_2(\xi_i), \ldots, \chi_n(\xi_i)$, abgesehen von der Reihenfolge, die n Wurzeln $\alpha_1, \alpha_2, \ldots, \alpha_n$ von $f(z) = 0$. Es gilt folgender Satz:

Man erhält sämtliche Permutationen der Galoisschen Gruppe $\mathfrak{G}$ der Gleichung $f(z) = 0$ und auch nur diese, sowie eine jede nur einmal, wenn man in der Permutation:

$$\begin{pmatrix} \chi_1(\xi_1) & \chi_2(\xi_1) & \cdots & \chi_n(\xi_1) \\ \chi_1(\xi_i) & \chi_2(\xi_i) & \cdots & \chi_n(\xi_i) \end{pmatrix}$$

die Größe ξ_i sämtliche g Wurzeln der Galoisschen Resolvente $G(\xi) = 0$ durchlaufen läßt.

Wählt man für $G(\xi)$ eine andere Galoissche Resolvente von $f(z) = 0$, so erscheinen die Permutationen der Galoisschen Gruppe $\mathfrak{G}$ von $f(z) = 0$ eventuell in anderer Reihenfolge.

Die Natur der Galoisschen Gruppe gibt über die Eigenschaften einer Gleichung folgende Aufschlüsse:

Eine Gleichung ist in dem der Betrachtung zugrunde liegenden Körper irreduzibel oder reduzibel, je nachdem ihre Galoissche Gruppe eine transitive oder intransitive Permutationsgruppe ist.

Eine irreduzible Gleichung ist in dem der Untersuchung zugrunde liegenden Körper P dann und nur dann primitiv (vgl. S. 210), *wenn ihre Galoissche Gruppe eine primitive Permutationsgruppe ist. Die Galoissche Gruppe jeder im Körper P irreduziblen imprimitiven Gleichung ist eine imprimitive Permutationsgruppe*

Die Galoissche Gruppe einer Normalgleichung und auch nur einer solchen Gleichung ist eine reguläre Permutationsgruppe, d. h. eine transitive Permutationsgruppe von derselben Ordnung wie Grad (vgl. S. 212).

Hat man eine Gleichung $f(z) = 0$ mit Koeffizienten aus einem Körper P und den Wurzeln $\alpha_1, \alpha_2, \ldots, \alpha_n$, so heißt eine ganze rationale Funktion $\varphi(\alpha_1, \alpha_2, \ldots, \alpha_n)$ von $\alpha_1, \alpha_2, \ldots, \alpha_n$ mit Koeffizienten aus P, deren Wert dem Körper P nicht angehört, *eine der Gleichung* $f(z) = 0$ *natürliche Irrationalität.* Eine solche natürliche Irrationalität kann auch als eine in dem durch die Galoissche Resolvente $G(\zeta) = 0$ bestimmten Galoisschen Körper (P, ζ_1) gelegene Größe definiert werden.

Eine natürliche Irrationalität φ bleibt nicht bei allen Permutationen der Galoisschen Gruppe $\mathfrak{G}$ der Gleichung $f(z) = 0$ numerisch ungeändert; denn sonst wäre sie bereits im Körper P gelegen.

Die Gesamtheit aller Permutationen der Galoisschen Gruppe $\mathfrak{G}$, *bei denen* $\varphi(\alpha_1, \alpha_2, \ldots, \alpha_n)$ *numerisch ungeändert bleibt, bildet eine Untergruppe* $\mathfrak{H}$ *von* $\mathfrak{G}$. *Die Gruppe* $\mathfrak{H}$ *heißt die zu der natürlichen Irrationalität* φ *zugehörige Gruppe. Umgekehrt gehören zu jeder Untergruppe* $\mathfrak{H}$ *der Galoisschen Gruppe* $\mathfrak{G}$ *eine und sogar unendlich viele natürliche Irrationalitäten* $\varphi(\alpha_1, \alpha_2, \ldots, \alpha_n)$, *die nur bei den Permutationen von* $\mathfrak{H}$ *numerisch ungeändert bleiben, bei allen* $\mathfrak{H}$ *nicht angehörigen Permutationen von* $\mathfrak{G}$ *ihren numerischen Wert ändern.*

Gehört die natürliche Irrationalität φ_1 zu der Untergruppe $\mathfrak{H}$ der Galoisschen Gruppe $\mathfrak{G}$ und ist $\mathfrak{G} = \mathfrak{H}G_1 + \mathfrak{H}G_2 + \cdots \mathfrak{H}G_\varrho$ die Zerlegung der Gruppe $\mathfrak{G}$ nach dem Modul $\mathfrak{H}$, so nimmt φ_1 den ϱ Nebengruppen entsprechend bei allen Permutationen von $\mathfrak{G}$ die ϱ numerisch verschiedenen Werte $\varphi_1, \varphi_2, \ldots, \varphi_\varrho$ an. *Diese genügen der im Körper* P *irreduziblen Gleichung*

$$\Phi(t) = (t - \varphi_1)(t - \varphi_2) \ldots (t - \varphi_\varrho) = 0.$$

Sie heißt eine *Resolvente von* $f(z) = 0$ und präziser, da alle ihre Wurzeln rationale Funktionen der Wurzeln von $f(z) = 0$ mit Koeffizienten aus P sind, eine *rationale Resolvente von* $f(z) = 0$. *Ist* $\mathfrak{J}$ *die invariante Untergruppe von* $\mathfrak{G}$, *welche die konjugierten Gruppen* $G_i^{-1}\mathfrak{H}G_i$ $(i = 1, 2, \ldots, \varrho)$ *gemein haben, so ist die Galoissche Gruppe von* $\Phi(t) = 0$ *mit der Faktorgruppe* $\mathfrak{G}/\mathfrak{J}$ *isomorph.* Die ϱ konjugierten Größen φ_i $(i = 1, 2, \ldots, \varrho)$ haben die Gruppen $G_i^{-1}\mathfrak{H}G_i$ $(i = 1, 2, \ldots, \varrho)$ als zugehörige Gruppen. Die Erweiterung des Körpers P durch eine Größe φ_i reduziert die

Galoissche Gruppe $\mathfrak{G}$ von $f(z) = 0$ auf die zu φ_i zugehörige Gruppe $G_i^{-1} \mathfrak{H} G_i$; die Adjunktion aller Wurzeln von $\Phi(t) = 0$ zum Körper P reduziert demnach die Galoissche Gruppe von $f(z) = 0$ auf die Gruppe $\mathfrak{J}$, den Durchschnitt aller konjugierten Gruppen.

Lagrangescher, von Galois erweiterter Satz (vgl. S. 221): *Ist φ_1 eine zur Untergruppe $\mathfrak{H}$ der Galoisschen Gruppe $\mathfrak{G}$ gehörige natürliche Irrationalität und bleibt eine zweite natürliche Irrationalität $\psi(\alpha_1, \alpha_2, \ldots, \alpha_n)$ bei allen Permutationen von $\mathfrak{H}$ numerisch ungeändert, so ist ψ eine rationale Funktion von φ_1 mit Koeffizienten aus* P. *Ist $\mathfrak{H}'$ die zu ψ zugehörige Untergruppe der Galoisschen Gruppe $\mathfrak{G}$ und $\sigma = \frac{h'}{h}$ der Index der Untergruppe $\mathfrak{H}$ von $\mathfrak{H}'$, so genügt φ_1 einer Gleichung σ^{ten} Grades, deren Koeffizienten dem durch Adjunktion von ψ erweiterten Körper* P *angehören; die Gleichung σ^{ten} Grades, der φ_1 genügt, ist in diesem erweiterten Körper irreduzibel.*

Korollar: Hat man zwei natürliche Irrationalitäten, die zu derselben Untergruppe der Galoisschen Gruppe gehören, so ist jede eine rationale Funktion der anderen mit Koeffizienten aus P. *Die rationalen Resolventen mit Koeffizienten aus* P, *denen sie genügen, gehen daher durch Tschirnhausentransformation auseinander hervor.*

Im Gegensatz zu den natürlichen Irrationalitäten spricht man von den *für die Gleichung $f(z) = 0$ akzessorischen Irrationalitäten. Hierunter versteht man jede dem Körper* P *nicht angehörige Größe, die sich nicht in die Form $\varphi(\alpha_1, \alpha_2, \ldots, \alpha_n)$ mit Koeffizienten aus* P *bringen läßt.* Die Adjunktion einer akzessorischen Irrationalität η zum Körper P ändert die Galoissche Gruppe $\mathfrak{G}$ von $f(z) = 0$ entweder überhaupt nicht oder reduziert sie auf eine ihrer Untergruppen $\mathfrak{H}$. *Reduziert die Adjunktion einer akzessorischen Irrationalität η die Galoissche Gruppe $\mathfrak{G}$ von $f(z) = 0$ auf eine Untergruppe $\mathfrak{H}$ und wählt man eine zu $\mathfrak{H}$ zugehörige natürliche Irrationalität φ, so ist φ eine rationale Funktion von η mit Koeffizienten aus* P. *Jede Größe η, deren Adjunktion die Galoissche Gruppe $\mathfrak{G}_g$ von $f(z) = 0$ auf $\mathfrak{H}_h$ reduziert, genügt einer in* P *irreduziblen Gleichung, deren Grad ein Vielfaches des Index $\frac{g}{h}$ der Gruppe $\mathfrak{H}_h$ von $\mathfrak{G}_g$ ist und nur dann gleich $\frac{g}{h}$ wird, wenn η keine akzessorische, sondern eine natürliche Irrationalität ist. In letzterem Fall ist auch η eine rationale Funktion von φ mit Koeffizienten aus* P.

Alle Wurzeln einer im Körper P *irreduziblen Gleichung, deren Adjunktion die Galoissche Gruppe der gegebenen Gleichung reduziert, sind entweder ausnahmslos natürliche oder ausnahmslos akzessorische Irrationalitäten.* Reduziert die Adjunktion einer Wurzel einer irreduziblen Gleichung die Galoissche Gruppe $\mathfrak{G}$ einer gegebenen Gleichung $f(z) = 0$, so reduziert sie auch die Adjunktion jeder anderen Wurzel der irreduziblen Gleichung, und zwar auf dieselbe oder eine innerhalb $\mathfrak{G}$ ähnliche Untergruppe.

Satz von C. Jordan (*Traité*, p. 269) und Hölder (*Math. Ann.* **34**, 47 (1889)): *Reduziert sich die Galoissche Gruppe* $\mathfrak{G}$ *einer Gleichung* $f(z) = 0$ *mit Koeffizienten aus* P *durch Adjunktion sämtlicher Wurzeln einer beliebigen (reduziblen oder irreduziblen) Gleichung* $F(t) = 0$ *mit Koeffizienten aus* P *auf eine Untergruppe* $\mathfrak{H}$, *so reduziert sich die Galoissche Gruppe* $\mathfrak{G}_1$ *von* $F(t) = 0$ *durch Adjunktion aller Wurzeln von* $f(z) = 0$ *auf eine Untergruppe* $\mathfrak{H}_1$. *Die Gruppen* $\mathfrak{H}$ *und* $\mathfrak{H}_1$ *sind invariante Untergruppen von* $\mathfrak{G}$ *bezw.* $\mathfrak{G}_1$, *ferner sind die Faktorgruppen* $\mathfrak{G}/\mathfrak{H}$ *und* $\mathfrak{G}_1/\mathfrak{H}_1$ *holoedrisch isomorph.*

Korollare: a) *Die Adjunktion einer einzigen Wurzel einer irreduziblen Hilfsgleichung* $F(t) = 0$ *reduziert die Galoissche Gruppe* $\mathfrak{G}$ *der Gleichung* $f(z) = 0$ *dann und nur dann auf eine invariante Untergruppe, wenn* $F(t) = 0$ *eine Normalgleichung im Körper* P *ist.*

b) Erniedrigt die Adjunktion aller Wurzeln einer Gleichung $F(t) = 0$ *mit einfacher Galoisscher Gruppe die Galoissche Gruppe von* $f(z) = 0$, *so zerfällt die Hilfsgleichung* $F(t) = 0$ *durch Adjunktion aller Wurzeln von* $f(z) = 0$ *in lauter Faktoren ersten Grades, d. h. die Wurzeln von* $F(t) = 0$ *sind rationale Funktionen derjenigen von* $f(z) = 0$ *mit Koeffizienten aus dem Körper* P *oder anders ausgedrückt: sie sind natürliche Irrationalitäten der Gleichung* $f(z) = 0$.

Da jede invariante Untergruppe einer primitiven Permutationsgruppe transitiv ist (vgl. S. 210), so folgt aus dem Jordan-Hölderschen Hauptsatz auch: *Eine in einem Körper* P *irreduzible primitive Gleichung, die durch Adjunktion aller Wurzeln einer Hilfsgleichung reduzibel wird, zerfällt in lauter Faktoren ersten Grades.*

Wird eine in einem Körper P *irreduzible imprimitive Gleichung durch Adjunktion aller Wurzeln einer Hilfsgleichung reduzibel, so zerfällt sie in lauter Faktoren des nämlichen Grades; diese haben in dem durch Adjunktion aller Wurzeln der Hilfsgleichung erweiterten Körper isomorphe Galoissche Gruppen* (vgl. die für imprimitive Gruppen geltenden Sätze auf S. 210, ferner Weber, *Algebra* **1**, 563).

Ist $f(z) = 0$ eine beliebige Gleichung mit Koeffizienten aus einem Körper P *und der Galoisschen Gruppe* $\mathfrak{G}$ *und besitzt diese die Kompositionsreihe* $\mathfrak{G}, \mathfrak{G}_1, \mathfrak{G}_2, \ldots, \mathfrak{G}_{\nu-1}, 1$ (vgl. S. 197), *so läßt sich die Behandlung der vorgelegten Gleichung auf eine Gleichungskette* $f_1(z) = 0, f_2(z) = 0, \ldots, f_\nu(z) = 0$ *von folgender Beschaffenheit zurückführen: Jede dieser Gleichungen ist in dem Körper* P, *nachdem man ihm sämtliche Wurzeln der voraufgehenden Gleichungen adjungiert hat, irreduzibel und besitzt eine einfache Gruppe, die mit den sukzessiven Faktorgruppen* $\mathfrak{G}/\mathfrak{G}_1$, $\mathfrak{G}_1/\mathfrak{G}_2, \ldots, \mathfrak{G}_{\nu-2}/\mathfrak{G}_{\nu-1}, \mathfrak{G}_{\nu-1}$ *holoedrisch isomorph ist. Die Galoissche Gruppe der vorgelegten Gleichung reduziert sich nach Adjunktion der Wurzeln der Hilfsgleichungen sukzessiv auf* $\mathfrak{G}_1, \mathfrak{G}_2, \ldots, \mathfrak{G}_{\nu-1}, 1$. *Infolge des voraufgehenden Korollars b) sind die Hilfsgleichungen in dem Körper, dem ihre Koeffizienten angehören, rationale Resolventen von* $f(z) = 0$. *Man kann die Hilfsgleichungen im besonderen stets so wählen, daß sie sämtlich Normalgleichungen sind, infolgedessen ihre Grade gleich den Ordnungen ihrer Galoisschen Gruppen werden und die Adjunktion je einer einzigen ihrer Wurzeln genügt.*

Für eine *veränderte Wahl der Hilfsgleichungen* gilt folgendes Theorem: Es seien $f_1(z) = 0, f_2(z) = 0, \ldots, f_\lambda(z) = 0$ irgendwelche Gleichungen ohne mehrfache Wurzeln von folgender Beschaffenheit: $f_1(z) = 0$ habe Koeffizienten aus P und eine einfache Galoissche Gruppe Γ_1. Die Gleichung $f_i(z) = 0$ habe Koeffizienten aus dem durch Adjunktion aller Wurzeln von $f_1(z) = 0, f_2(z) = 0, \ldots, f_{i-1}(z) = 0$ erweiterten Körper P und in diesem Oberkörper, dem ihre Koeffizienten angehören, eine einfache Galoissche Gruppe Γ_i. Reduziert sich die Galoissche Gruppe $\mathfrak{G}$ von $f(z) = 0$ durch Adjunktion aller Wurzeln von

$$f_1(z) = 0, f_2(z) = 0, \ldots, f_\lambda(z) = 0$$

auf die Einheit, d. h. drücken sich die Wurzeln von $f(z) = 0$ rational durch die von $f_1(z) = 0, f_2(z) = 0, \ldots, f_\lambda(z) = 0$ mit Koeffizienten aus P aus, so kommen unter den λ Gruppen $\Gamma_1, \Gamma_2, \ldots, \Gamma_\lambda$ ν Gruppen vor, die mit den Faktorgruppen

$$\mathfrak{G}/\mathfrak{G}_1, \mathfrak{G}_1/\mathfrak{G}_2, \ldots, \mathfrak{G}_{\nu-2}/\mathfrak{G}_{\nu-1}, \mathfrak{G}_{\nu-1},$$

vielleicht nur in anderer Reihenfolge, holoedrisch isomorph sind. *Durch Herbeiziehen beliebiger Hilfsgleichungen, deren Wurzeln akzessorische Irrationalitäten sind, wird niemals das Auftreten der mit den Faktorgruppen* $\mathfrak{G}/\mathfrak{G}_1, \mathfrak{G}_1/\mathfrak{G}_2, \ldots, \mathfrak{G}_{\nu-2}/\mathfrak{G}_{\nu-1}, \mathfrak{G}_{\nu-1}$ *isomorphen Gruppen vermieden: diese müssen stets, nur möglicher-*

weise in verschiedener Reihenfolge, auftreten. Ist $\nu = \lambda$, so sind alle irreduziblen Faktoren von $f_1(z) = 0, f_2(z) = 0, \ldots, f_\nu(z) = 0$ rationale Resolventen von $f(z) = 0$ (Hölder, *Math. Ann.* **34**, 49 (1889)).

Unter einer *allgemeinen Gleichung*

$$a_0 z^n + a_1 z^{n-1} + \cdots + a_n = 0$$

vom n^{ten} Grade versteht man eine solche, bei der die Koeffizienten $a_0, a_1, \ldots, a_n$ willkürliche Variablen sind. Der Körper P, in dem die allgemeine Gleichung n^{ten} Grades betrachtet wird, muß alle rationalen Funktionen von $a_0, a_1, a_2, \ldots, a_n$ enthalten. Die Koeffizienten dieser Funktionen können entweder ausschließlich als gewöhnliche rationale Zahlen angenommen oder als Zahlen eines beliebigen Zahlkörpers oder noch allgemeiner als irgendwelche Zahlen vorausgesetzt werden. Im letzteren Fall enthält der Körper P außer den Unbestimmten $a_0, a_1, a_2, \ldots, a_n$ alle numerischen Größen. *Die Galoissche Gruppe der allgemeinen Gleichung n^{ten} Grades ist die symmetrische Gruppe n^{ten} Grades der Ordnung $n!$. In dem durch Adjunktion der Quadratwurzel aus der Diskriminante erweiterten Körper reduziert sich die Galoissche Gruppe auf die alternierende Gruppe n^{ten} Grades der Ordnung $\frac{n!}{2}$, die für $n > 4$ einfach ist. Für $n > 4$ ist die allgemeine Gleichung n^{ten} Grades nicht mehr algebraisch auflösbar.* (Vgl. S. 252.) Für die Fälle $n = 4$ und $n = 3$ vgl. § 8.

Ist $f(z) = 0$ die *allgemeine* Gleichung n^{ten} Grades, so sind ihre Wurzeln ebenso wie die Koeffizienten als willkürliche Variablen anzusehen. Eine zu einer Untergruppe $\mathfrak{H}_h$ der Galoisschen Gruppe der Ordnung $n!$ zugehörige natürliche Irrationalität $\varphi(\alpha_1, \alpha_2, \ldots, \alpha_n)$ ist alsdann nichts anderes als eine zu der Gruppe $\mathfrak{H}_h$ zugehörige rationale Funktion $\varphi(\alpha_1, \alpha_2, \ldots, \alpha_n)$ der n Variablen $\alpha_1, \alpha_2, \ldots, \alpha_n$, die bei allen Permutationen von $\mathfrak{H}_h$ und keinen weiteren formal ungeändert bleibt. Nimmt die Funktion φ bei den Permutationen der symmetrischen Gruppe die ϱ Werte $\varphi_1 = \varphi, \varphi_2, \ldots, \varphi_\varrho$ an, so ist die Resolvente

$$\Phi(t) = (t - \varphi_1)(t - \varphi_2) \ldots (t - \varphi_\varrho) = 0$$

mit der auf S. 280 definierten transformierten Gleichung identisch. Auf diese Weise ordnen sich die Betrachtungen auf S. 217 und S. 221 der allgemeinen Galoisschen Theorie unter. Im besondern geht aus dem von Galois verallgemeinerten Lagrangeschen Satz (vgl. S. 301) der Lagrangesche Satz (vgl. S. 221) hervor.

Eine Gleichung n^{ten} Grades, deren Galoissche Gruppe die symmetrische Gruppe n^{ten} Grades ist, heißt nach Kronecker (*Festschrift*, § 12, *Journ. f. Math.* **92** (1882), *Ges. Werke* **2**, 287) *eine Gleichung ohne Affekt.* Daß nicht nur die allgemeine Gleichung n^{ten} Grades affektlos ist, sondern auch spezielle Gleichungen mit numerischen Koeffizienten existieren, die ebenfalls affektlos sind, besagt der Satz von Hilbert (*Journ. f. Math.* **110**, 128 (1892), Weber, *Algebra* **1**, 652, E. Maillet, *Journ. de math.* (5) **5**, 205 (1899), M. Bauer, *Journ. f. Math.* **132**, 33 (1907)): *In jedem beliebigen, durch eine algebraische Zahl bestimmten Körper kann man stets Gleichungen von gegebenem Grade n konstruieren, deren Koeffizienten gewöhnliche rationale Zahlen sind und deren Galoissche Gruppe in eben jenem Körper (also a fortiori im absoluten Rationalitätsbereich) die symmetrische Gruppe n^{ten} Grades ist, und ebenso solche, deren Galoissche Gruppe in jenem Körper die alternierende Gruppe n^{ten} Grades ist.*

Ist $n > 4$, so gibt es für eine Gleichung n^{ten} Grades, deren Galoissche Gruppe die symmetrische Gruppe ist, abgesehen von Resolventen zweiten Grades, keine rationalen Resolventen von niedrigerem als n^{tem} Grade. Diese sind, abgesehen von $n = 6$, Tschirnhausenresolventen der ursprünglichen Gleichung. (Vgl. das Bertrandsche Problem auf S. 216.) Für $n = 4$ gibt es rationale Resolventen dritten Grades, vgl. die Gleichungen (3′) und (3″) auf S. 287 und 288.

Außer den im § 1 genannten Lehrbüchern, dem Enzyklopädieartikel von Hölder und den in diesem Paragraphen zitierten Schriften vgl. man für die Galoissche Theorie noch die Darstellungen von Bachmann, *Math. Ann.* **18**, 449 (1881), Bolza, *Am. J. math.* **13**, 59 (1891), *Math. Ann.* **42**, 253 (1893), Pierpont, *Annals of math.* (2) **1**, 113 u. **2**, 22 (1900), Mertens, *Sitzungsb. d. Wiener Akad.* **111**, Abt. IIa, 17 (1902).

§ 11. Abelsche Gleichungen.

Eine Gleichung $f(z) = 0$ mit lauter verschiedenen Wurzeln und mit Koeffizienten aus einem Körper P heißt eine *Abelsche Gleichung* (präziser eine im Körper P Abelsche Gleichung), wenn sich alle ihre Wurzeln als rationale Funktionen einer einzigen, z. B. α_1 mit Koeffizienten aus P ausdrücken lassen und außerdem stets

$$\Theta_i(\Theta_k(\alpha_1)) = \Theta_k(\Theta_i(\alpha_1))$$

ist, wenn $\Theta_i(\alpha_1)$ und $\Theta_k(\alpha_1)$ zwei beliebige Wurzeln von $f(x) = 0$ sind. $\Theta_i(\alpha_1)$ und $\Theta_k(\alpha_1)$ bedeuten rationale Funktionen von α_1 mit Koeffizienten aus P.

Die Galoissche Gruppe einer Abelschen Gleichung ist eine Abelsche Gruppe (vgl. S. 202). *Jede irreduzible Gleichung, deren Galoissche Gruppe eine Abelsche Gruppe ist, ist eine Abelsche Gleichung.* Läßt man den Zusatz irreduzibel fort, so gilt nur das Theorem: Jeder irreduzible Faktor der linken Seite einer Gleichung mit kommutativer Galoisscher Gruppe ist eine Abelsche Gleichung. Eine reduzible Gleichung mit kommutativer Galoisscher Gruppe braucht nämlich nicht so beschaffen zu sein, daß sich alle ihre Wurzeln rational durch eine von ihnen ausdrücken lassen. C. Jordan, *Traité*, p. 287 bezeichnet jede Gleichung mit kommutativer Galoisscher Gruppe als eine Abelsche Gleichung. Die oben gegebene engere Definition ist die ursprüngliche Abels (§ 4 seiner fundamentalen Arbeit, vgl. oben S. 253).

Eine Gleichung n^{ten} Grades mit n verschiedenen Wurzeln und mit Koeffizienten aus einem Körper P heißt eine *einfache Abelsche Gleichung*, wenn sich ihre Wurzeln in der Form

$$\alpha_1, \quad \alpha_2 = \Theta(\alpha_1), \quad \alpha_3 = \Theta(\alpha_2) = \Theta^2(\alpha_1), \ldots$$
$$\alpha_n = \Theta(\alpha_{n-1}) = \Theta^{n-1}(\alpha_1), \; \Theta(\alpha_n) = \Theta^n(\alpha_1) = \alpha_1$$

ausdrücken lassen, wobei $\Theta(\alpha_1)$ eine rationale Funktion von α_1 mit Koeffizienten aus P bedeutet. *Die Galoissche Gruppe einer einfachen Abelschen Gleichung n^{ten} Grades mit den n Wurzeln $\alpha_1, \alpha_2, \ldots, \alpha_n$ besteht entweder aus einer zyklischen Permutation $C = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ und ihren Potenzen oder aus den Permutationen $1, C^e, C^{2e}, \ldots, C^{(f-1)e}$, wobei $n = ef$ ist. Im ersten Fall ist die Gleichung irreduzibel, im zweiten Fall zerfällt ihre linke Seite in e irreduzible einfache Abelsche Gleichungen des Grades f.*

Umgekehrt: *Eine Gleichung n^{ten} Grades, deren Galoissche Gruppe aus der durch die zyklische Permutation $C = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ und ihren Potenzen erzeugten Gruppe oder einer ihrer Untergruppen besteht, ist eine einfache Abelsche Gruppe.*

Eine rationale Funktion von n beliebigen Größen $\alpha_1, \alpha_2, \ldots, \alpha_n$, die bei Anwendung der zyklischen Permutation $C = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ formal ungeändert bleibt, heißt eine *zyklische Funktion der n Größen* $\alpha_1, \alpha_2, \ldots, \alpha_n$.

Aus den beiden letzten Sätzen folgt: *Eine Gleichung n^{ten} Grades mit n verschiedenen Wurzeln und Koeffizienten aus einem Körper* P *ist dann und nur dann eine einfache Abelsche Glei-*

chung, wenn sich ihre Wurzeln in eine solche Reihenfolge bringen lassen, daß ihre zyklischen Funktionen mit Koeffizienten aus P *dem Körper* P *angehören.* Die einfachen Abelschen Gleichungen werden daher auch als *zyklische Gleichungen* bezeichnet.

Ist ε eine beliebige n^{te} Einheitswurzel (vgl. § 12), so ist

$$(\alpha_1 + \varepsilon\alpha_2 + \varepsilon^2\alpha_3 + \cdots + \varepsilon^{n-1}\alpha_n)^n$$

im Körper der n^{ten} Einheitswurzeln eine zyklische Funktion der n Größen $\alpha_1, \alpha_2, \ldots, \alpha_n$. Für eine einfache Abelsche Gleichung mit Koeffizienten aus P und Wurzeln $\alpha_1, \alpha_2, \ldots, \alpha_n$ ist demnach diese Größe eine in dem durch Adjunktion der n^{ten} Einheitswurzeln erweiterten Körper P gelegene bekannte Zahl b. Die für den erweiterten Körper natürliche Irrationalität

$$\alpha_1 + \varepsilon\alpha_2 + \varepsilon^2\alpha_3 + \cdots + \varepsilon^{n-1}\alpha_n$$

der Gleichung $f(z) = 0$ wird als ihre *Lagrangesche Resolvente*[1]) (Lagrange, *Œuvres* 3, 331) bezeichnet. Man findet sie durch Ziehen der n^{ten} Wurzel aus b. Wählt man für ε eine *primitive n^{te} Einheitswurzel* von der Beschaffenheit, daß b nicht verschwindet, so lassen sich die Wurzeln einer einfachen Abelschen Gleichung in dem durch Adjunktion der n^{ten} Einheitswurzeln erweiterten Körper durch Ausziehen einer einzigen n^{ten} Wurzel, nämlich derjenigen aus der Größe b, finden (Abel, a. a. O., § 3). Ist $n = p^\lambda$ eine Primzahlpotenz, so kann man stets eine derartige primitive $p^{\lambda\,\text{te}}$ Einheitswurzel angeben, daß b von Null verschieden ist. Wegen der für $b = 0$ vorzunehmenden Modifikation vgl. Weber, *Algebra* **1**, 589. Stets gilt der Satz: *Alle Wurzeln einer im Körper* P *einfachen Abelschen Gleichung lassen sich bestimmen, indem man dem Körper* P *eine primitive n^{te} Einheitswurzel adjungiert und außerdem eine n^{te} Wurzel aus einer einzigen Größe zieht.*

Sind zwei Wurzeln einer irreduziblen Gleichung vom Primzahlgrad in einem derartigen Verhältnis, daß man die eine rational durch die andere ausdrücken kann, so ist die Gleichung eine irreduzible einfache Abelsche Gleichung. Normalgleichung vom Primzahlgrad und irreduzible einfache Abelsche Gleichung sind identische Begriffe (Abel, a. a. O., § 3).

1) Unter Umständen nennt man auch eine Wurzel einer Resolventengleichung eine Resolvente, daher der Name „Lagrangesche Resolvente". Diese genügt der Gleichung $z^n = b$, die in dem durch Adjunktion von ε erweiterten Körper $(\mathsf{P}, \varepsilon)$ eine rationale Resolvente der vorgelegten einfachen Abelschen Gleichung im Sinne der Bezeichnung auf S. 300 ist.

Eine einfache Abelsche Gleichung mit *reellen Koeffizienten*, bei welcher der zugrunde liegende *Körper* P *nur aus lauter reellen Zahlen* besteht, hat entweder nur reelle oder nur imaginäre Wurzeln. Der erste Fall tritt stets ein, wenn die Gleichung ungeraden Grades ist. Zur Lösung einer einfachen Abelschen Gleichung n^{ten} Grades mit zugrunde liegendem *reellem* Körper P ist die Adjunktion einer primitiven n^{ten} Einheitswurzel, das Ausziehen einer Quadratwurzel aus einer positiven Zahl und die Teilung eines Winkels in n gleiche Teile erforderlich (Abel, a. a. O. § 3). Die Vereinfachung gegenüber den einfachen Abelschen Gleichungen, bei denen der Körper P auch imaginäre Zahlen enthalten kann, liegt darin, daß bei jenen außer der Kenntnis einer primitiven n^{ten} Einheitswurzel noch eine n^{te} Wurzel aus einer Größe der Form $\varrho(\cos\varphi + i\sin\varphi)$ zu ziehen ist und dies das Ausziehen einer n^{ten} Wurzel aus der positiven Größe ϱ und die Teilung eines Winkels in n gleiche Teile erfordert.

Da jeder irreduzible Faktor einer Abelschen Gleichung wiederum eine Abelsche Gleichung ist, kann man sich auf die Betrachtung der irreduziblen Abelschen Gleichungen beschränken. Jede irreduzible (auch nicht einfache) Abelsche Gleichung ist eine Normalgleichung; daher ist die Ordnung ihrer Galoisschen Gruppe gleich dem Grade n der Gleichung. Ist $\mathfrak{G}$ die Galoissche Gruppe einer irreduziblen Abelschen Gleichung des Grades n, so kann man für die Abelsche Gruppe $\mathfrak{G}$ ein System von σ erzeugenden Basiselementen $A_1, A_2, \ldots, A_\sigma$ der Ordnungen $g_1, g_2, \ldots, g_\sigma$ derartig auswählen, daß das Produkt $g_1 g_2 \ldots g_\sigma = n$ wird (vgl. S. 202). Wählt man alsdann σ zu den durch

$$(A_2, A_3, \ldots, A_\sigma), \quad (A_1, A_3, \ldots, A_\sigma), \ldots, (A_1, A_2, \ldots, A_{\sigma-1})$$

erzeugten Gruppen zugehörige natürliche Irrationalitäten, so genügt jede von ihnen einer irreduziblen *einfachen* Abelschen Gleichung des Grades g_1 bzw. $g_2, \ldots, g_\sigma$, wie aus den Ergebnissen auf S. 300 folgt. Hieraus ergibt sich der Satz: *Die Wurzeln jeder beliebigen irreduziblen Abelschen Gleichung n^{ten} Grades mit Koeffizienten aus* P *lassen sich als rationale Funktionen mit Koeffizienten aus* P *von Wurzeln irreduzibler einfacher Abelscher Gleichungen mit Koeffizienten aus* P *darstellen, bei denen das Produkt der Grade gleich n ist.*

Da man die Basis einer Abelschen Gruppe (vgl. S. 203) verschieden wählen kann, so ergeben sich folgende Sätze:

I. Die Auflösung einer irreduziblen Abelschen Gleichung des Grades n mit der Galoisschen Gruppe $\mathfrak{G}$ kann auf die Auflösung einer Reihe nebeneinander zu lösender irreduzibler einfacher Abelscher Gleichungen der Grade $p_1^{\tau_1}, p_2^{\tau_2}, \ldots, p_\sigma^{\tau_\sigma}$ mit Koeffizienten aus dem der Betrachtung zugrunde liegenden Körper reduziert werden; hierbei ist $n = p_1^{\tau_1} p_2^{\tau_2} \ldots p_\sigma^{\tau_\sigma}$, und $p_1, p_2, \ldots, p_\sigma$ bedeuten Primzahlen, die unter Umständen nicht alle voneinander verschieden sind und deren Anzahl alsdann größer als die Zahl der in n enthaltenen verschiedenen Primzahlen ist.

II. Die Auflösung einer irreduziblen Abelschen Gleichung des Grades n läßt sich auf die Auflösung einer Reihe nebeneinander zu lösender, irreduzibler einfacher Abelscher Gleichungen mit Koeffizienten aus dem der Betrachtung zugrunde liegenden Körper reduzieren, so daß der Grad einer jeden Gleichung dieser Reihe teilbar durch den Grad der folgenden Gleichung oder ihm gleich wird und das Produkt der Grade aller Gleichungen gleich der Zahl n ist. Die Anzahl der Gleichungen ist gleich dem Range der Gruppe (vgl. S. 203).

Die Behandlung einer beliebigen Abelschen Gleichung ist demnach auf diejenige einfacher Abelscher Gleichungen zurückführbar. Da diese zur Lösung nur eine Wurzelausziehung und Kenntnis der Einheitswurzeln erfordern, die sich ebenfalls durch Radikale darstellen lassen, so folgt: *Jede Abelsche Gleichung ist algebraisch auflösbar* (Abel, a. a. O. § 4).

Die Wurzeln jeder im Körper P *irreduziblen Abelschen Gleichung n^{ten} Grades lassen sich mit Hilfe einer Wurzel α der Gleichung in die Form*

$$(A) \qquad \Theta_1^{h_1} \Theta_2^{h_2} \ldots \Theta_\sigma^{h_\sigma}(\alpha)$$

bringen, so daß in dieser Form jede Wurzel der Gleichung einmal und auch nur einmal erscheint, falls h_i die Zahlen $0, 1, 2, \ldots, g_i - 1$ durchläuft. Die mit einander vertauschbaren Θ_i bedeuten die Operationen einer rationalen Funktion mit Koeffizienten aus P *und die Zahl h_i die h_i-fache Wiederholung ($i = 1, 2, \ldots, \sigma$); g_i ist die kleinste Zahl, für die $\Theta_i^{g_i}(\alpha) = 1$ ist.* Nennt man $g_1, g_2, \ldots, g_\sigma$ *die Wiederholungszahlen,* so gilt folgender Satz: *Jeder Basis der Galoisschen Gruppe $\mathfrak{G}$ einer irreduziblen Abelschen Gleichung entspricht eine Darstellung der Gleichungswurzeln in der Form (A), so daß die Wiederholungszahlen mit den Ordnungszahlen der Basiselemente übereinstimmen.*

Jede irreduzible Abelsche Gleichung, deren Grad keine

Primzahl ist, ist als Normalgleichung imprimitiv, da ihre Galoissche Gruppe als reguläre Gruppe imprimitiv ist (vgl. S. 212). *Die Auflösung einer irreduziblen Abelschen Gleichung vom Primzahlpotenzgrad p^τ kann auf die von τ nacheinander zu lösenden irreduziblen Abelschen Gleichungen vom Primzahlgrad p reduziert werden.*

Satz von Kronecker (*Monatsb. d. Berl. Akad.* (1853), 373, ebenda (1877), 849, Weber, *Algebra* **2**, 762, Hilbert, *Math.-Ver.* **4**, *Die Theorie der algebraischen Zahlkörper*, Kap. 23, Mertens, *Journ. f. Math.* **131**, 87 (1906), Weber, ebenda **132**, 167 (1907), *Math. Ann.* **67**, 32 (1909)): *Alle Wurzeln Abelscher Gleichungen mit ganzzahligen Koeffizienten sind rationale Funktionen von Einheitswurzeln mit rationalen Zahlenkoeffizienten, und alle rationalen Funktionen von Einheitswurzeln mit rationalen Zahlenkoeffizienten sind Wurzeln ganzzahliger Abelscher Gleichungen.* Anders ausgedrückt: *Jeder durch eine irreduzible Abelsche Gleichung aus dem natürlichen Rationalitätsbereich hervorgehende Zahlkörper ist ein Kreisteilungskörper, d. h. ein durch eine Einheitswurzel $e^{\frac{2\pi i}{m}}$ erzeugter Körper oder ein in diesem Körper enthaltener Unterkörper, und jeder Kreisteilungskörper ist ein Abelscher Körper.*

Ebenso wie nach dem voraufgehenden Satz die Exponentialfunktion $e^{2\pi i\omega}$ für rationale Werte des Arguments ω mit den Abelschen Gleichungen mit ganzzahligen Koeffizienten verknüpft ist, trifft dies auch für die *elliptische Modulfunktion, z. B. die Invariante $j(\omega)$, zu, wenn ω Wurzel einer ganzzahligen quadratischen Gleichung mit negativer Diskriminante ist.* Eine derartige Modulfunktion heißt nach Kronecker *singulär. Eine singuläre Invariante $j(\omega)$ genügt einer irreduziblen Gleichung mit ganzzahligen Koeffizienten, der sogenannten Klassengleichung, die auch noch in dem durch Adjunktion der Quadratwurzel aus der Diskriminante der quadratischen Gleichung erweiterten Körper irreduzibel bleibt und dann eine irreduzible Abelsche Gleichung ist.* Daß die singulären elliptischen Modulfunktionen durch Wurzelzeichen ausdrückbar sind, was aus der Natur der Klassengleichung, eine „relativ Abelsche" Gleichung zu sein, folgt, ahnte bereits Abel, vgl. die Bemerkungen von Sylow in Abels *Œuvres* **2**, 316; den ersten Beweis gab Kronecker, *Monatsb. d. Berl. Akad.* (1857), 455, (1862), 363 und (1877), 851. Eine zusammenfassende Darstellung mit zahlreichen Literaturnachweisen findet man in Webers Artikel *„Komplexe*

Multiplikation", Enzyklopädie der math. Wiss. **1**, 716 und in dem dritten Buche des dritten Bandes seiner Algebra. Vgl. auch F. Klein, *Ausgewählte Kapitel der Zahlentheorie, Autographierte Vorlesungen*, Göttingen 1896. Es gilt auch umgekehrt folgender Satz, Kroneckers „liebster Jugendtraum" (Brief von Kronecker an Dedekind, veröff. von Frobenius, *Sitzungsb. d. Berl. Akad.* (1895), 115): *Alle in einem imaginären quadratischen Körper Abelschen Gleichungen lassen sich durch die Körper der singulären Moduln und durch die Kreisteilungskörper lösen.* Einen Beweis hat zuerst Fueter, *Journ. f. Math.* **132**, 255 (1907) geliefert; vgl. auch Fueter, *Göttinger Diss.* (1903). Eine Ausdehnung dieser Untersuchungen ist der von Hilbert, (*Gött. Nachr.* (1898), abgedr. *Acta math.* **26**, 99 (1902)) und Furtwängler (*Math. Ann.* **63**, 1 (1907)) stammende Nachweis der Existenz eines Klassenkörpers für einen beliebigen algebraischen Grundkörper P, der im einfachsten Fall der obige imaginäre quadratische Körper ist. Dieser allgemeine Klassenkörper wird durch eine im Körper P Abelsche Gleichung mit Koeffizienten aus P oder aus einem in diesem Körper enthaltenen Unterkörper definiert.

§ 12. Die Einheitswurzeln. Die Kreisteilungsgleichung. Gaußsche Summen. Reine Gleichung.

Alle Wurzeln der Gleichung $z^n = 1$ *heißen* n^{te} *Einheitswurzeln; sie sind sämtlich voneinander verschieden und werden, wenn* k *die Werte* $0, 1, 2, \ldots, n-1$ *durchläuft, in der Form* $\varepsilon_k = \cos\frac{2k\pi}{n} + i\sin\frac{2k\pi}{n}$ *gegeben.*

Die Bilder der Größen ε_k in der sog. Gaußschen Ebene der komplexen Zahlen liegen auf einem Kreise mit dem Radius Eins und bilden ein reguläres n-Eck, dessen erster Eckpunkt sich auf der Achse der reellen Zahlen befindet; die Teilung des Kreises in n gleiche Teile hängt demnach von der Lösung der Gleichung $z^n = 1$ ab.

Befriedigt ε *die Gleichung* $z^n = 1$, *so ist auch* ε^m *eine* n^{te} *Einheitswurzel, wenn* m *irgendeine ganze positive oder negative Zahl einschließlich der Null ist. Das Produkt zweier beliebiger* n^{ter} *Einheitswurzeln ist wiederum eine* n^{te} *Einheitswurzel.*

Ist ε *zugleich* n^{te} *und* m^{te} *Einheitswurzel, so ist es auch* d^{te} *Einheitswurzel, wenn* d *der größte gemeinsame Teiler von* n *und* m *ist.*

Die Summe s_k *der* k^{ten} *Potenzen aller* n^{ten} *Einheitswurzeln ist*

Null, außer wenn k durch n teilbar ist. Ist k durch n teilbar, so ist $s_k = n$.

Eine n^{te} Einheitswurzel heißt eine primitive n^{te} Einheitswurzel, wenn ε keiner Gleichung $z^m = 1$ genügt, wobei $m < n$ ist. Eine nicht primitive n^{te} Einheitswurzel heißt eine *imprimitive n^{te} Einheitswurzel.*

Eine charakteristische Eigenschaft einer primitiven n^{ten} Einheitswurzel ε ist, daß ihre Potenzen $\varepsilon, \varepsilon^2, \varepsilon^3, \ldots, \varepsilon^n$ sämtliche verschiedene n^{te} Einheitswurzeln ergeben. Durch Potenzieren einer imprimitiven n^{ten} Einheitswurzel erhält man niemals sämtliche n^{te} Einheitswurzeln.

Jede imprimitive n^{te} Einheitswurzel genügt einer Gleichung $z^m = 1$, wobei m ein Divisor von n ist. Umgekehrt ist jede Wurzel von $z^m = 1$, wenn m ein Teiler von n ist, eine imprimitive n^{te} Einheitswurzel. Jede imprimitive n^{te} Einheitswurzel ist primitive m^{te} Einheitswurzel, wobei m ein gewisser Divisor von n ist.

Die Gleichung $z^n = 1$ hat soviel primitive n^{te} Einheitswurzeln, wie die Anzahl $\varphi(n)$ derjenigen ganzen positiven Zahlen beträgt, die kleiner als n und relativ prim zu n sind. Ist p eine Primzahl, so ist jede Wurzel der Gleichung $z^p = 1$, abgesehen von der Zahl 1, primitive p^{te} Einheitswurzel. Es ist $\varphi(p) = p - 1$.

Ist p^k die k^{te} Potenz der Primzahl p, so findet man alle Wurzeln der Gleichung $z^{p^k} = 1$, indem man je eine Wurzel $\omega_1, \omega_2, \ldots, \omega_k$ der k Gleichungen $z^p = 1, z^p = \omega_1, z^p = \omega_2, \ldots, z^p = \omega_{k-1}$ wählt und das Produkt $\omega_1 . \omega_2 \ldots \omega_k$ bildet. Dann und nur dann, wenn ω_1 eine primitive p^{te} Einheitswurzel ist, d. h. $\omega_1 \neq 1$, ist das Produkt $\omega_1 . \omega_2 \ldots \omega_k$ eine primitive $p^{k\,te}$ Einheitswurzel. Es ist $\varphi(p^k) = p^{k-1}(p-1) = p^k\left(1 - \frac{1}{p}\right)$.

Ist $n = p_1^{k_1} \cdot p_2^{k_2} \ldots p_l^{k_l}$, wobei $p_1, p_2, \ldots, p_l$ lauter untereinander verschiedene Primzahlen sind, so findet man alle Wurzeln der Gleichung $z^n = 1$, indem man je eine Wurzel $\lambda_1, \lambda_2, \ldots, \lambda_l$ der Gleichungen $z^{p_1^{k_1}} = 1, z^{p_2^{k_2}} = 1, \ldots, z^{p_l^{k_l}} = 1$ wählt und ihr Produkt $\lambda_1 . \lambda_2 \ldots \lambda_l$ bildet. Dieses Produkt stellt dann und nur dann eine primitive n^{te} Einheitswurzel dar, wenn alle Größen $\lambda_1, \lambda_2, \ldots, \lambda_l$ primitive $p_1^{k_1}$ bzw. $p_2^{k_2}, \ldots, p_l^{k_l te}$ Einheitswurzeln sind. Es ist $\varphi(n) = \varphi(p_1^{k_1})\,\varphi(p_2^{k_2}) \ldots \varphi(p_l^{k_l}) =$

$$= p_1^{k_1} \cdot p_2^{k_2} \ldots p_l^{k_l}\left(1 - \frac{1}{p_1}\right)\left(1 - \frac{1}{p_2}\right) \cdots \left(1 - \frac{1}{p_l}\right).$$

Ist ε eine primitive n^{te} Einheitswurzel, so sind, wenn k zu n

relativ prim ist, alle k^{ten} Potenzen von ε und auch nur diese Potenzen primitive n^{te} Einheitswurzeln.

Die Summe aller primitiven n^{ten} Einheitswurzeln wird mit $\mu(n)$ bezeichnet. Ist n durch das Quadrat einer Primzahl teilbar, so hat die zahlentheoretische Funktion $\mu(n)$ den Wert 0, ist n das Produkt von t verschiedenen Primzahlen, so hat $\mu(n)$ den Wert $(-1)^t$; $\mu(1)$ ist gleich $+1$. (Moebius, *Journ. f. Math.* **9**, 109 (1832), *Ges. Werke* **4**, 595, Mertens, *Journ. f. Math.* **77**, 289 (1874)).

Die $\varphi(n)$ primitiven n^{ten} Einheitswurzeln genügen der Gleichung $\varphi(n)^{ten}$ Grades:

$$Z_n = \Pi\left(z^{\frac{n}{t}} - 1\right)^{\mu(t)} = 0;$$

hierbei ist das Produkt über alle Teiler t von n, einschließlich n und 1, zu erstrecken, und $\mu(t)$ bedeutet die oben definierte zahlentheoretische Funktion. Die Gleichung $Z_n = 0$ hat ganzzahlige Koeffizienten; sie ist im absoluten Rationalitätsbereiche irreduzibel, und bleibt es auch noch bei Adjunktion einer primitiven q^{ten} Einheitswurzel, falls n und q relativ prime Zahlen sind. (Kronecker, *Journ. de math.* **19**, 192 (1854), *Ges. Werke* **1**, 92, *Journ. f. Math.* **100**, 79 (1887)).

Da die sukzessiven Potenzen jeder Wurzel der Gleichung $Z_n = 0$ alle n Wurzeln der Gleichung $z^n = 1$ liefern, von der die Teilung des Kreises in n gleiche Teile abhängt, heißt die Gleichung $Z_n = 0$ die *Kreisteilungsgleichung.*

Für eine Primzahl $n = p$ lautet die Kreisteilungsgleichung:

$$\frac{z^p - 1}{z - 1} = z^{p-1} + z^{p-2} + \cdots + 1 = 0.$$

Für eine Primzahlpotenz $n = p^k$ lautet die Kreisteilungsgleichung

$$\frac{z^{p^k} - 1}{z^{p^{k-1}} - 1} = z^{(p-1)p^{k-1}} + z^{(p-2)p^{k-1}} + z^{(p-3)p^{k-1}} + \cdots + z^{p^{k-1}} + 1 = 0.$$

Ist $n = p_1^{k_1} \cdot p_2^{k_2} \ldots p_l^{k_l}$ ($p_1, p_2, \ldots, p_l$ verschiedene Primzahlen), so ist

$$Z_n = \Pi\left(z^{\frac{n}{t}} - 1\right)^{\mu(t)} = \frac{\left(z^n - 1\right)\left(z^{\frac{n}{p_1 p_2}} - 1\right)\left(z^{\frac{n}{p_1 p_3}} - 1\right)\ldots}{\left(z^{\frac{n}{p_1}} - 1\right)\left(z^{\frac{n}{p_2}} - 1\right)\left(z^{\frac{n}{p_3}} - 1\right)\ldots} = 0.$$

Die Diskriminante der Gleichung $Z_n = 0$ hat den Wert $\frac{(-1)^{\frac{\varphi(n)}{2}} \cdot n^{\varphi(n)}}{p_1^{\frac{\varphi(n)}{\varphi(p_1)}} \cdot p_2^{\frac{\varphi(n)}{\varphi(p_2)}} \cdots p_l^{\frac{\varphi(n)}{\varphi(p_l)}}}$, wobei φ die oben eingeführte zahlentheoretische Funktion bedeutet. (Vgl. die elementare Ableitung von Rados, *Journ. f. Math.* **131**, 49 (1906)).

Die Irreduzibilität der Gleichung $Z_p = 0$ im absoluten Rationalitätsbereich hat zuerst Gauß (*Disquisitiones arithmeticae* (1801), Art. 341) bewiesen. Andere Beweise stammen von Kronecker, *Journ. f. Math.* **29**, 280 (1845), *Journ. de math.* (2) **1**, 399 (1856), *Ges. Werke* **1**, 1 und 99, Schoenemann und Eisenstein (Schoenemann, *Journ. f. Math.* **32**, 100 (1846), Eisenstein, ebenda **39**, 166 (1850), Schoenemanns Reklamation, ebenda **40**, 188 (1850), a. a. O. Zeile 8 v. u. ist **32**, § 61 zu lesen). Die Irreduzibilität von $Z_p^k = 0$ im absoluten Rationalitätsbereich bei Serret, *Journ. de math.* **15**, 296 (1850). Daß die Gleichung $Z_n = 0$ für *beliebiges n* im absoluten Rationalitätsbereich irreduzibel ist, und es auch sogar bleibt, wenn eine Wurzel einer ganzzahligen irreduziblen Gleichung adjungiert wird, deren Diskriminante zu n relativ prim ist, hat Kronecker (*Journ. de math.* **19**, 177 (1854), *Ges. Werk* **1**, 75, *Journ. f. Math.* **100**, 79 (1887)) zuerst bewiesen. Andere Beweise der Irreduzibilität von $Z_n = 0$ für beliebiges n stammen von Dedekind, *Journ. f. Math.* **54**, 27 (1857), Arndt, *Journ. f. Math.* **56**, 178 (1859), Lebesgue, *Journ. de math.* (2) **4**, 105 (1859).

Die Kreisteilungsgleichung $Z_n = 0$ ist eine irreduzible Abelsche Gleichung. Ihre Galoissche Gruppe ist eine transitive Abelsche Gruppe des Grades und der Ordnung $\varphi(n)$, die holoedrisch isomorph ist mit der Gruppe der Zahlklassen der zu n relativ primen Zahlen (vgl. S. 203, Z. 10 v. u.). *Die Gruppe ist dann und nur dann zyklisch, wenn n die Potenz einer ungeraden Primzahl, das Doppelte hiervon oder 2 oder 4 ist. Für diese und nur für diese Werte der Zahl n ist die Kreisteilungsgleichung eine irreduzible einfache Abelsche Gleichung.*

Aus dem Satz: „Eine Normalgleichung ist dann und nur dann durch *Quadratwurzeln* lösbar, wenn ihr Grad eine Potenz von 2 ist", folgt: *Die Wurzeln der Kreisteilungsgleichung $Z_n = 0$ sind dann und nur dann durch Quadratwurzeln darstellbar, wenn der Grad $\varphi(n)$ von $Z_n = 0$ eine Potenz von 2 ist.* Da sich alle Wurzeln von $z^n = 1$ durch Potenzieren einer beliebigen

primitiven n^{ten} Einheitswurzel ergeben, folgt *das Theorem von* Gauß (*Disquisitiones arithmeticae*, Art. 365, 366): *Damit die Gleichung* $z^n = 1$ *durch Quadratwurzeln lösbar sein soll, ist notwendig und hinreichend, daß* n *entweder gleich einer Primzahl der Form* $2^h + 1$ *oder gleich* 2^m *oder gleich einem Produkt mehrerer verschiedener Zahlen dieser Formen ist. Dann und nur dann, wenn sich die Gleichung* $z^n = 1$ *durch Quadratwurzeln auflösen läßt, kann der Kreisumfang mit Zirkel und Lineal in* n *gleiche Teile geteilt werden.* Für dieselben Werte von n kann auch der *Umfang der Lemniskate* mit Hilfe von Zirkel und Lineal in n gleiche Teile geteilt werden (Gauß, *Disquisitiones arithmeticae*, Art. 335, Abel, *Recherches sur les fonctions elliptiques, Œuvres* **1**, p. 361, Eisenstein, *Journ. f. Math.* **39**, 160, 224, 275 (1850), Schwering, *Journ. f. Math.* **107**, 196 (1890)). Eine Zahl der Form $2^h + 1$ kann nur dann Primzahl sein, wenn $h = 2^k$ ist; jedoch sind nicht alle Zahlen der Form $2^{2^k} + 1$ Primzahlen. $k = 0, 1, 2, 3, 4$ liefert die Primzahlen 3, 5, 17, 257, 65537; hingegen erhält man für $k = 5$ die Zahl $2^{2^5} + 1 = 4294967297$, die durch 641 teilbar ist. Vermutlich existiert nur eine endliche Anzahl von Primzahlen der Form $2^{2^k} + 1$. Die Gleichung $z^3 = 1$ hat die Wurzeln

$$1, -\frac{1}{2} \pm \frac{i}{2}\sqrt{3},$$

die Gleichung $z^5 = 1$ besitzt die primitive Wurzel

$$\cos\frac{2\pi}{5} + i\sin\frac{2\pi}{5} = \frac{1}{4}(-1 + \sqrt{5} + i\sqrt{10 + 2\sqrt{5}}),$$

die Wurzeln der Gleichung $z^{17} = 1$ findet man in Gauß, *Disquisitiones arithmeticae*, Art. 354. Die Gleichung $z^{257} = 1$ hat Richelot, *Journ. f. Math.* **9**, 1 (1832) behandelt. Für die Gleichung $z^{65537} = 1$ vgl. Hermes, *Gött. Nachr.* (1894), 170. Unterhalb 100 gibt es im ganzen 25 Zahlen, für die $z^n = 1$ durch Quadratwurzeln lösbar ist, nämlich $n = 2, 3, 4, 5, 6, 8, 10, 12, 15, 16, 17, 20, 24, 30, 32, 34, 40, 48, 51, 60, 64, 68, 80, 85, 96$.

Die geometrische Konstruktion des regulären Siebzehnecks mittelst des Lineals und eines festen Kreises geben v. Staudt, *Journ. f. Math.* **24**, 251 (1842) und Schroeter, ebenda **75**, 13 (1873), nach Mascheroni unter alleiniger Anwendung des Zirkels Gérard, *Math. Ann.* **48**, 390 (1897). Die Konstruktion des regulären 257-Ecks bei Pascal, *Rend. Accad. di Na-*

poli (1887). Für nicht mittelst Zirkels und Lineals durchführbare Fälle der Kreisteilung, nämlich 7, 13 und 97 vgl. Pascal, *Giorn. di mat.* **25**, 82 (1887), für 19 und 37 Amaldi, ebenda **30**, 141 (1892). Für die hier behandelten Fragen vgl. besonders Klein, *Vorträge über ausgewählte Fragen der Elementargeometrie*, ausg. von Taegert, Leipzig 1895, und Enriques, *Fragen der Elementargeometrie*, Teil II, deutsch von Fleischer, Leipzig 1907.

Wir beschränken uns darauf, von den Fällen, in denen die Kreisteilungsgleichung eine irreduzible einfache Abelsche Gleichung ist ($n = 2, 4, p^\lambda, 2p^\lambda$, vgl. oben), denjenigen näher zu betrachten, für den n eine Primzahl ist.

Ist p eine Primzahl, so lassen sich die Wurzeln der Gleichung

$$\frac{z^p - 1}{z - 1} = z^{p-1} + z^{p-2} + \cdots + 1 = 0$$

in der Form $\alpha, \Theta(\alpha), \Theta^2(\alpha), \ldots, \Theta^{p-2}(\alpha)$ anordnen, hierbei ist α eine beliebige Wurzel, $\Theta^i(\alpha) = \alpha^{g^i}$, und g bedeutet eine primitive Wurzel für die Primzahl p.[1])

Da die Gleichung $\frac{z^p - 1}{z - 1} = 0$ eine im Körper der reellen Zahlen einfache Abelsche Gleichung ist, kann ihre Lösung bewerkstelligt werden (vgl. S. 308) *mittelst einer primitiven $(p-1)^{ten}$ Einheitswurzel, Teilung eines Winkels in $p - 1$ gleiche Teile und Ausziehung einer Quadratwurzel aus einer positiven Zahl, die hier die Zahl p ist* (Gauß, *Disquisitiones arithmeticae*, Art. 360).

Zur weiteren Behandlung von $\frac{z^p - 1}{z - 1} = 0$ hat Gauß, *Disquisitiones arithmeticae*, Art. 343, die *f-gliedrigen Perioden* eingeführt. Sei $p - 1$ irgendwie in zwei ganzzahlige positive Faktoren e und f zerlegt. Ist α eine beliebige Wurzel der Gleichung $\frac{z^p - 1}{z - 1} = 0$, so heißen die e Funktionen:

$$\eta_\lambda = \alpha^{g^\lambda} + \alpha^{g^{e+\lambda}} + \alpha^{g^{2e+\lambda}} + \cdots \alpha^{g^{(f-1)e+\lambda}} \qquad (\lambda = 0, 1, 2, \ldots, e-1)$$

die e Gaußschen f-gliedrigen Perioden. Die e Perioden sind Wurzeln der Gleichung e^{ten} Grades:

$$\Phi = (\eta - \eta_0)(\eta - \eta_1)(\eta - \eta_2) \cdots (\eta - \eta_{e-1}) = 0.$$

---

1) g heißt eine primitive Wurzel für die Primzahl p, wenn die Zahlen $g, g^2, g^3, \ldots, g^{p-1}$ bei ihrer Division durch p, von der Reihenfolge abgesehen, die Zahlen $1, 2, \ldots, p-1$ ergeben.

Diese hat ganzzahlige Koeffizienten und ist eine im absoluten Rationalitätsbereich irreduzible einfache Abelsche Gleichung. Jede der Perioden ist also eine ganze rationale Funktion jeder anderen mit rationalen Zahlenkoeffizienten. Nach Adjunktion irgendeiner Wurzel dieser Normalgleichung zerfällt die Gleichung $\frac{z^p-1}{z-1}=0$ in das Produkt der e Gleichungen:

$$\chi_\lambda = \left(z-\alpha^{g^\lambda}\right)\left(z-\alpha^{g^{e+\lambda}}\right)\left(z-\alpha^{g^{2e+\lambda}}\right)\ldots\left(z-\alpha^{g^{(f-1)e+\lambda}}\right)=0.$$
$$(\lambda=0,1,2,\ldots,e-1)$$

Die Koeffizienten jeder dieser e Gleichungen $\chi_\lambda=0$ sind lineare ganze Funktionen der e Perioden $\eta_0,\eta_1,\ldots,\eta_{e-1}$ mit ganzzahligen Koeffizienten, und jede dieser Gleichungen ist in dem durch Adjunktion der e Perioden erweiterten Körper eine irreduzible einfache Abelsche Gleichung. Läßt sich die Zahl f irgendwie in zwei ganzzahlige Faktoren e' und f' spalten, so bilde man die sog. f'-gliedrigen Perioden:

$$\eta_\lambda' = \alpha^{g^{\lambda e}} + \alpha^{g^{(e'+\lambda)e}} + \alpha^{g^{(2e'+\lambda)e}} + \cdots + \alpha^{g^{((f'-1)e'+\lambda)e}}$$
$$(\lambda=0,1,2,\ldots,e'-1).$$

Die Größen $\eta_0', \eta_1', \ldots, \eta_{e'-1}'$ genügen einer Gleichung, deren Koeffizienten lineare ganze Funktionen von $\eta_0,\eta_1,\ldots,\eta_{e-1}$ mit ganzzahligen Koeffizienten sind. Diese Gleichung ist in dem durch die Perioden $\eta_0,\eta_1,\ldots,\eta_{e-1}$ erweiterten Körper eine irreduzible Abelsche Gleichung. Nach Adjunktion einer Wurzel dieser Normalgleichung zerfällt jede der e Gleichungen χ_λ in e' Gleichungen vom Grade f'. Jede dieser Gleichungen ist in dem durch $\eta_0', \eta_1', \ldots, \eta_{e'-1}'$ erweiterten Körper eine irreduzible einfache Abelsche Gleichung. Dieses Verfahren läßt sich fortsetzen, indem man f' in Faktoren spaltet. Vgl. Gauß, *Disquisitiones arithmeticae*, Art. 342—354.

Die *Galoissche Theorie* liefert den Kern dieses Verfahrens. Die Wurzeln der Gleichung $\frac{z^p-1}{z-1}=0$ seien mit

$$\alpha_i = \Theta^i(\alpha) = \alpha^{g^i} \qquad (i=0,1,2,\ldots,p-2)$$

bezeichnet. Die Galoissche Gruppe $\mathfrak{G}$ der Gleichung $\frac{z^p-1}{z-1}=0$ wird von der zyklischen Permutation $C=(\alpha,\alpha_1,\alpha_2,\ldots,\alpha_{p-2})$ und ihren Potenzen gebildet. Die Permutation C^e und ihre Potenzen bilden eine Untergruppe $\mathfrak{G}_1$ von $\mathfrak{G}$ des Index e. Jede der e Perioden $\eta_0,\eta_1,\eta_2,\ldots,\eta_{e-1}$ ist eine zu $\mathfrak{G}_1$ zugehörige

natürliche Irrationalität. Die e Perioden vertauschen sich bei allen $\mathfrak{G}_1$ nicht angehörigen Permutationen von $\mathfrak{G}$ nur untereinander und genügen daher einer irreduziblen Normalgleichung $\Phi = 0$, deren Galoissche Gruppe zu $\mathfrak{G}/\mathfrak{G}_1$ isomorph ist und aus der zyklischen Permutation $T = (\eta_0, \eta_1, \eta_2, \ldots, \eta_{e-1})$ und ihren Potenzen besteht. Adjungiert man η_0 zum absoluten Rationalitätsbereich, so reduziert sich die Galoissche Gruppe $\mathfrak{G}$ unserer Gleichung $\frac{z^p - 1}{z - 1} = 0$ auf die intransitive Gruppe $\mathfrak{G}_1$ der Ordnung f. Ist $f = e'f'$, so erzeugt die Permutation $C^{e e'}$ mit ihren Potenzen eine Untergruppe $\mathfrak{G}_2$ von $\mathfrak{G}_1$ vom Index e'; die Funktionen $\eta_0', \eta_1', \ldots, \eta_{e'-1}'$ sind zu $\mathfrak{G}_2$ zugehörige natürliche Irrationalitäten, die sich bei allen $\mathfrak{G}_2$ nicht angehörigen Permutationen von $\mathfrak{G}_1$ nur untereinander permutieren.

Für eine ungerade Primzahl p kann man $p - 1$ in das Produkt $2 \cdot \frac{(p-1)}{2}$ zerlegen. Für $e = 2$, $f = \frac{p-1}{2}$ wird:

$$\eta_0 = \alpha + \alpha^{g^2} + \alpha^{g^4} + \cdots + \alpha^{g^{p-3}},$$
$$\eta_1 = \alpha^g + \alpha^{g^3} + \cdots + \alpha^{g^{p-2}}.$$

Die zwei Größen η_0 und η_1 genügen der quadratischen Gleichung $\Phi = \eta^2 + \eta + \frac{1 - (-1)^{\frac{p-1}{2}} p}{4} = 0$ mit ganzzahligen Koeffizienten. Es wird:

$$\eta_0 = -\tfrac{1}{2} \pm \tfrac{1}{2}\sqrt{(-1)^{\frac{p-1}{2}} \cdot p}, \quad \eta_1 = -\tfrac{1}{2} \mp \tfrac{1}{2}\sqrt{(-1)^{\frac{p-1}{2}} \cdot p},$$

$$\eta_0 - \eta_1 = \pm \sqrt{(-1)^{\frac{p-1}{2}} \cdot p}.$$

Durch Adjunktion von $\sqrt{(-1)^{\frac{p-1}{2}} \cdot p}$ zum absoluten Rationalitätsbereich wird $\frac{z^p - 1}{z - 1} = 0$ reduzibel und zerfällt in zwei Faktoren des Grades $\frac{p-1}{2}$.

Bei η_0 sind die Exponenten von α die quadratischen Reste, bei η_1 die quadratischen Nichtreste unter den Zahlen $1, 2, \ldots, p-1$ nach p. Daher ist: $\eta_0 - \eta_1 = \sum_{s=1}^{s=p-1} \left(\frac{s}{p}\right) \alpha^s$, wobei $\left(\frac{s}{p}\right)$ das Legendresche Symbol ist, das den Wert $+1$

oder -1 hat, je nachdem s quadratischer Rest oder Nichtrest der Primzahl p ist.

Ein berühmtes Problem, dem Gauß (1811) (*Ges. Werke* 2, 9, deutsche Ausg. von E. Netto in *Die sechs Beweise des Fundamentaltheorems über quadratische Reste von Gauß* in Ostwalds *Klassikern der exakt. Wiss.* Nr. 122, S. 51) eine besondere Abhandlung gewidmet hat, ist die *Bestimmung des Vorzeichens* der oben auf der rechten Seite von $\eta_0 - \eta_1$ auftretenden Quadratwurzel, die davon abhängt, welche Einheitswurzel man für α wählt. Ist $\alpha = e^{\frac{2\pi i \mu}{p}}$, wobei μ zu p relativ prim ist, und demnach $\eta_0 - \eta_1 = \sum_{s=1}^{s=p-1} \left(\frac{s}{p}\right) e^{\frac{2\pi i \mu s}{p}}$, so hat die Summe $\sum_{s=1}^{s=p-1} \left(\frac{s}{p}\right) e^{\frac{2\pi i \mu s}{p}}$, *die man als Gaußsche Summe bezeichnet, den Wert* $+\left(\frac{\mu}{p}\right) i^{\left(\frac{p-1}{2}\right)^2} \sqrt{p}$; *hierbei ist* $\left(\frac{\mu}{p}\right)$ *das Legendresche Symbol und* $i = \sqrt{-1}$. Mithin ergibt sich für $p \equiv 1 \pmod 4$ die Wertbestimmung $\left(\frac{\mu}{p}\right) \cdot \sqrt{p}$ und für $p \equiv 3 \pmod 4$ gleich $i\left(\frac{\mu}{p}\right)\sqrt{p}$. Die Gleichung

$$\sum_{s=1}^{s=p-1} \left(\frac{s}{p}\right) e^{\frac{2\pi i \mu s}{p}} = \left(\frac{\mu}{p}\right) i^{\left(\frac{p-1}{2}\right)^2} \sqrt{p}$$

bleibt noch richtig, wenn p eine beliebige ganze, aber keine Primzahl mehr ist und $\left(\frac{s}{p}\right)$ das Jacobische Symbol bedeutet, das jedesmal, wenn s und p einen gemeinsamen Teiler besitzen, den Wert 0 zu erhalten hat.[1]) $\left(\frac{\mu}{p}\right)$ ist dann ebenfalls das Jacobische Symbol. Auch können μ und p einen Teiler gemein haben; in diesem Fall ist $\left(\frac{\mu}{p}\right) = 0$ zu nehmen. Eine Darstellung der in die Zahlentheorie gehörigen Gaußschen Summen findet man bei Bachmann, *Analytische Zahlentheorie,* Leipzig 1894, S. 145.

---

1) Sind s und p zwei relativ prime Zahlen und ist p in die Primzahlfaktoren $p = p_1 p_2 p_3 \ldots$ zerlegbar, so ist das Jacobische Symbol definiert durch

$$\left(\frac{s}{p}\right) = \left(\frac{s}{p_1}\right)\left(\frac{s}{p_2}\right)\left(\frac{s}{p_3}\right)\ldots, \text{ wobei } \left(\frac{s}{p_1}\right), \left(\frac{s}{p_2}\right), \left(\frac{s}{p_3}\right), \ldots$$

die Legendreschen Symbole sind.

Weitere Angaben über Einheitswurzeln findet man in der Monographie von Bachmann, *Die Lehre von der Kreisteilung*, Leipzig 1872, Weber, *Algebra* **1**, 596ff., **2**, 69 u. 736 und D. Hilbert, *Die Theorie der algebraischen Zahlkörper, Math.-Ver.* **4** (1897), S. 325 (der Kreiskörper). Vgl. auch § 11, S. 310.

Wir schließen diesen Paragraphen mit der Betrachtung der *reinen Gleichung* $z^n = a$, die auch als *binomische Gleichung* bezeichnet wird. *Ist n eine Primzahl p, so ist die Gleichung $z^p - a = 0$ in jedem Körper* P, *der a enthält, dann und nur dann reduzibel, wenn a die p^{te} Potenz einer Zahl aus* P *ist.* Ist also die Gleichung $z^p = a$ reduzibel, so hat sie eine dem Körper P angehörige Wurzel. (Abel, *Œuvres* **1**, 72, Jordan, *Traité*, p. 300, Mertens, *Monatsh. f. Math.* **2**, 291 (1891), Kneser, *Journ. f. Math.* **106**, 48 (1890)).

Für beliebiges n gilt folgender Satz (Capelli, *Rend. Acc. di Napoli*, Dez. 1897, Febr. 1898, *Math. Ann.* **54**, 602 (1901), Vahlen, *Acta math.* **19**, 195 (1895), Wendt, *Math. Ann.* **53**, 450 (1900)): *Die Gleichung $z^n - a = 0$ ist in einem beliebigen Körper* P, *der a enthält, dann und nur dann reduzibel, wenn a die h^{te} Potenz einer Zahl aus* P *ist, wobei h ein Teiler von n ist, oder wenn n durch 4 teilbar und $-a$ das Vierfache der vierten Potenz einer Zahl aus* P *ist.*

Ist α_0 eine Wurzel der Gleichung $z^n - a = 0$, so ergeben sich ihre sämtlichen Wurzeln in der Form $\alpha_i = \varepsilon^i \alpha_0$ $(i = 0, 1, 2, \ldots, n-1)$, wobei ε eine primitive n^{te} Einheitswurzel ist. Jede reine Gleichung $z^n - a = 0$ ist in einem Körper, der die n^{ten} Einheitswurzeln enthält, eine Abelsche Gleichung. In dem kleinsten Körper, der die Größe a enthält[1]*, ist die Gleichung $z^n - a = 0$ so beschaffen, daß sich alle ihre Wurzeln als rationale Funktionen von zwei geeigneten α_0, α_1 in der Form $\alpha_i = \left(\frac{\alpha_1}{\alpha_0}\right)^i \alpha_0$ $(i = 0, 1, 2, \ldots, n-1)$ ausdrücken lassen.*

§ 13. Algebraisch auflösbare Gleichungen.

Eine Gleichung $f(z) = 0$ mit Koeffizienten aus einem Körper P heißt *algebraisch auflösbar* (präziser eine im Körper P algebraisch auflösbare Gleichung), wenn sich ihre sämtlichen Wurzeln aus den Größen des Körpers P durch eine endliche Anzahl von Wurzelausziehungen finden lassen. Man sieht sofort, daß die Exponenten bei sämtlichen Wurzelausziehungen als Prim-

1) Also ohne Adjunktion von Einheitswurzeln.

zahlen angenommen werden können. Weber, *Algebra* **1**, 647 nennt die auflösbaren Gleichungen, wie man kürzer statt algebraisch auflösbar sagt, *metazyklisch.*

Die n Wurzeln $\alpha_1, \alpha_2, \ldots, \alpha_n$ einer auflösbaren Gleichung lassen sich stets in die Form: $\alpha_i = r_i(\varphi_\lambda, \varphi_{\lambda-1}, \ldots, \varphi_2, \varphi_1)$ $(i = 1, 2, \ldots, n)$ bringen, wobei die n Funktionen r_i ganze rationale Funktionen ihrer Argumente $\varphi_1, \varphi_2, \ldots, \varphi_\lambda$ mit Koeffizienten aus dem der Untersuchung zugrunde liegenden Körper P *bedeuten. Die Größen $\varphi_1, \varphi_2, \ldots, \varphi_\lambda$ bestimmen sich hierbei nacheinander durch eine Kette reiner Gleichungen: $\varphi_1^{\tau_1} = R_1$, $\varphi_2^{\tau_2} = R_2(\varphi_1)$,*

$$\varphi_3^{\tau_3} = R_3(\varphi_1, \varphi_2), \ldots, \quad \varphi_\lambda^{\tau_\lambda} = R_\lambda(\varphi_1, \varphi_2, \ldots, \varphi_{\lambda-1}).$$

Die Zahlen $\tau_1, \tau_2, \ldots, \tau_\lambda$ bedeuten Primzahlen; die Größen R_i $(i = 2, \ldots, \lambda)$ sind ganze rationale Funktionen von $\varphi_1, \varphi_2, \ldots, \varphi_{i-1}$ mit Koeffizienten aus P, *und R_1 ist eine dem Körper* P *angehörige Zahl.* Man kann es stets so einrichten, daß die Größen $\varphi_1, \varphi_2, \ldots, \varphi_\lambda$, d. h. die auftretenden Radikale, ganze rationale Funktionen der n Wurzeln von $f(z) = 0$ sind, und die Koeffizienten dieser Funktionen dem durch Adjunktion von Einheitswurzeln erweiterten Körper P angehören. *Die Größen $\varphi_1, \varphi_2, \ldots, \varphi_\lambda$ sind also bei Zugrundelegung des Körpers* P *akzessorische Irrationalitäten, die aber in dem durch Adjunktion von Einheitswurzeln erweiterten Körper natürliche Irrationalitäten sind.* Abel hat in seiner grundlegenden Arbeit (vgl. oben S. 252, *Œuvres* **1**, 75) bei der Untersuchung der Form, die man den Radikalen φ bei algebraisch auflösbaren Gleichungen geben kann, das Auftreten der Einheitswurzeln nicht betont; dies geschah erst durch Jordan (*Traité,* p. 270) und Kronecker (*Monatsb. d. Berl. Akad.* (1879), 206). Daß die bei der Lösung der Gleichungen *dritten* und *vierten* Grades auftretenden Radikale in die Form rationaler Funktionen der Gleichungswurzeln und Einheitswurzeln gebracht werden können, geht bereits aus Lagranges Angaben in seinen Réflexions hervor. (Vgl. S. 283 und 287.)

Da die reinen Gleichungen sich auf Abelsche Gleichungen zurückführen lassen, folgt:

Damit eine Gleichung auflösbar sei, ist notwendig und hinreichend, daß ihre Behandlung auf eine Kette irreduzibler Abelscher Gleichungen vom Primzahlgrad reduziert werden kann. Diese Behandlungsweise verdankt man Jordan, *Traité,* p. 386. Vgl. hierzu auch Bendixson, *Acta math.* **27**, 317 (1903).

Eine Gleichung ist dann und nur dann auflösbar, wenn ihre Galoissche Gruppe eine auflösbare Gruppe ist. (Vgl. S. 199.)

Hat man eine im Körper P auflösbare Gleichung $f(z) = 0$ mit der Galoisschen Gruppe $\mathfrak{G}$ der Ordnung g, deren Kompositionsreihe $\mathfrak{G}, \mathfrak{G}_1, \mathfrak{G}_2, \ldots, \mathfrak{G}_{f-1}, 1$ lautet und deren Indexreihe $\frac{g}{g_1}, \frac{g_1}{g_2}, \ldots, \frac{g_{f-2}}{g_{f-1}}, g_{f-1}$ infolge der Auflösbarkeit aus lauter Primzahlen besteht, so gestaltet sich die *Behandlung der auflösbaren Gleichung* $f(z) = 0$ *auf Grund der Galoisschen Theorie* folgendermaßen: Man bestimmt eine natürliche Irrationalität η_1 der Wurzeln von $f(z) = 0$, die zur Gruppe $\mathfrak{G}_1$ gehört, ebenso eine natürliche Irrationalität η_2, die zur Gruppe $\mathfrak{G}_2$ gehört, usw., schließlich eine zur Gruppe $\mathfrak{G}_{f-1}$ gehörige natürliche Irrationalität η_{f-1} der Wurzeln von $f(z) = 0$ und eine zur identischen Permutation 1 gehörige natürliche Irrationalität η_f. Dann ist η_1 eine rationale Funktion von η_2 mit Koeffizienten aus P, ebenso ist η_2 eine rationale Funktion von η_3 mit Koeffizienten aus P usw., schließlich η_{f-1} eine rationale Funktion von η_f mit Koeffizienten aus P. Die Größe η_1 genügt einer im Körper P irreduziblen einfachen Abelschen Gleichung vom Primzahlgrade $\frac{g}{g_1}$, deren Galoissche Gruppe mit der Faktorgruppe $\mathfrak{G}/\mathfrak{G}_1$ holoedrisch isomorph ist. Nach Adjunktion von η_1 genügt die Größe η_2 einer in dem erweiterten Körper (P, η_1) irreduziblen einfachen Abelschen Gleichung vom Primzahlgrade $\frac{g_1}{g_2}$, deren Galoissche Gruppe mit der Faktorgruppe $\mathfrak{G}_1/\mathfrak{G}_2$ holoedrisch isomorph ist. Nach Adjunktion von η_2 genügt die Größe η_3 einer irreduziblen einfachen Abelschen Gleichung vom Primzahlgrad $\frac{g_2}{g_3}$, deren Galoissche Gruppe mit der Faktorgruppe $\mathfrak{G}_2/\mathfrak{G}_3$ holoedrisch isomorph ist usw., schließlich genügt η_f in dem durch Adjunktion von η_{f-1} erweiterten Körper einer irreduziblen einfachen Abelschen Gleichung des Grades g_{f-1}, deren Galoissche Gruppe mit der Gruppe $\mathfrak{G}_{f-1}$ holoedrisch isomorph ist. Die Wurzeln der vorgelegten Gleichung $f(z) = 0$ sind ganze rationale Funktionen von η_f mit Koeffizienten aus P. Dieses Verfahren reduziert, wie es dem Jordanschen Satz entspricht, die Auflösung der Gleichung $f(z) = 0$ auf eine Kette von nacheinander zu lösenden einfachen Abelschen Gleichungen mit Gruppen von Primzahlordnung, die mit den Faktorgruppen $\mathfrak{G}/\mathfrak{G}_1, \mathfrak{G}_1/\mathfrak{G}_2, \ldots, \mathfrak{G}_{f-1}$ holoedrisch isomorph sind. Bei einer Reduktion auf Gleichungen mit einfachen Gruppen sind diese Gruppen nicht zu vermeiden, sie können eventuell nur in anderer

Reihenfolge auftreten. *Im Gegensatz zu den bei der Reduktion auf eine Kette von reinen Gleichungen oben* (S. 321) *eingeführten* φ *sind die Größen* $\eta_1, \eta_2, \ldots, \eta_f$ *natürliche Irrationalitäten des Körpers* P, *ohne daß man diesem Einheitswurzeln zu adjungieren braucht.*

Läßt sich auch nur eine Wurzel einer irreduziblen algebraischen Gleichung durch Wurzelausziehungen finden, so trifft dies für alle Wurzeln zu, und die Gleichung ist algebraisch auflösbar. (Abel, *Œuvres* **2**, 221, vgl. auch Selivanoff, *Acta math.* **19**, 88 (1895).)

Aus dem Satz von Galois (vgl. oben S. 210) und der Tatsache, daß die Ordnung jeder transitiven Permutationsgruppe durch ihren Grad teilbar ist, folgt der Satz von Abel (*Œuvres* **2**, 222 und 262): *Jede auflösbare irreduzible Gleichung, deren Grad sich durch wenigstens zwei verschiedene Primzahlen teilen läßt, ist imprimitiv. Ist also* n *durch zwei verschiedene Primzahlen teilbar, so zerfällt die vorgelegte Gleichung* n^{ten} *Grades durch Adjunktion der Wurzeln einer auflösbaren Gleichung* s^{ten} *Grades in Faktoren* r^{ten} *Grades, wobei* $n = rs$. *Die Behandlung auflösbarer Gleichungen läßt sich also auf das Studium der auflösbaren primitiven Gleichungen vom Primzahlpotenzgrad zurückführen.* (Vgl. Galois, *Œuvres*, p. 51.)

Hat man eine irreduzible primitive auflösbare Gleichung des Primzahlpotenzgrades p^k, so lassen sich ihre Wurzeln stets in der Form $\alpha_{\xi_1' \xi_2' \ldots \xi_k'}$ derartig anordnen, daß $\xi_1', \xi_2', \ldots, \xi_k'$ die Zahlen $0, 1, 2, \ldots, p-1$ durchlaufen und alle Permutationen der Galoisschen Gruppe der Gleichung das Aussehen $\begin{pmatrix} \alpha_{\xi_1' \xi_2' \cdots \xi_k'} \\ \alpha_{\xi_1 \xi_2 \cdots \xi_k} \end{pmatrix}$ haben; die Indices $\xi_1, \xi_2, \ldots, \xi_k$ ergeben sich hierbei aus $\xi_1', \xi_2', \ldots, \xi_k'$ durch Substitutionen der Form

$$\xi_i = a_{i1}\xi_1' + a_{i2}\xi_2' + \cdots + a_{ik}\xi_k' + a_{i,k+1}, \quad (i = 1, 2, \cdots, k)$$

wobei a_{ij} $(i = 1, 2, \ldots, k;\ j = 1, 2, \ldots, k+1)$ ganze Zahlen mod p mit nicht verschwindender Determinante $|a_{ij}|$ $(i, j = 1, 2, \ldots, k)$ bedeuten.

Die Galoissche Gruppe einer irreduziblen primitiven auflösbaren Gleichung des Primzahlpotenzgrades p^k *ist also entweder mit der allgemeinen vollen linearen Kongruenzgruppe der Ordnung* $(p^k - 1)(p^k - p)(p^k - p^2) \cdots (p^k - p^{k-1})p^k$ *in* k *Variablen mit ganzzahligen Koeffizienten mod* p *oder mit einer ihrer Untergruppen isomorph.* (Vgl. S. 247.) Für $k = 1$ erhält man die von Kronecker sog. *metazyklische Gruppe* des Grades p und der Ordnung $p(p-1)$, die stets auflösbar ist. (Vgl. S. 213,

Z. 21 auf S. 213 lies $|z, \alpha + \beta z|$ statt $|z, \alpha z + \beta|$). Hieraus ergeben sich die für $k = 1$ gültigen Galoisschen Theoreme:

Eine irreduzible Gleichung vom Primzahlgrad ist dann und nur dann auflösbar, wenn ihre Wurzeln sich so schreiben lassen, daß ihre Galoissche Gruppe die metazyklische Gruppe oder eine ihrer transitiven Untergruppen ist. (Galois, *Œuvres*, p. 48.)

Eine irreduzible Gleichung vom Primzahlgrad ist dann und nur dann auflösbar, wenn alle ihre Wurzeln rationale Funktionen zweier beliebiger Wurzeln mit Koeffizienten aus dem der Betrachtung zugrunde liegenden Körper sind. (Galois, *Œuvres*, p. 49.)

Hat eine irreduzible Gleichung vom Primzahlgrad mit reellen Koeffizienten mehr als eine reelle Wurzel, so ist sie bei Zugrundelegung eines reellen Körpers sicher nicht auflösbar, wenn sie nicht lauter reelle Wurzeln hat.

Aus der Natur der metazyklischen Gruppe und ihrer Untergruppen (vgl. S. 213) folgt:

Jede im Körper P *irreduzible auflösbare Gleichung vom Primzahlgrade p kann mittelst zweier irreduzibler einfacher Abelscher Gleichungen des Grades $p - 1$ bezw. $\frac{p-1}{\delta}$ und des Grades p gelöst werden.* Die Wurzeln einer im Körper P irreduziblen auflösbaren Gleichung p^{ten} Grades lassen sich nämlich so anordnen, daß zwischen ihnen die Beziehungen stattfinden $\alpha_1 = \Theta(\alpha_0)$, $\alpha_2 = \Theta(\alpha_1)$, ..., $\alpha_0 = \Theta(\alpha_{p-1})$. Die Operation Θ ist hierbei eine ganze rationale Funktion mit Koeffizienten aus einem Körper (P, φ); dieser geht aus dem der Betrachtung zugrunde liegenden Körper P durch Adjunktion einer Wurzel φ einer im Körper P irreduziblen einfachen Abelschen Gleichung des Grades $p - 1$ oder $\frac{p-1}{\delta}$ hervor, wobei δ ein Teiler von $p - 1$ ist. (Kronecker, *Monatsb. d. Berl. Akad.* (1853), 369.)

Als Endziel der Behandlung algebraisch auflösbarer Gleichungen sah bereits Abel (*Œuvres* 2, 222ff., 266) an, *Ausdrücke zu konstruieren,* die aus dem gegebenen Körper P durch Wurzelziehungen hervorgehen und die Doppeleigenschaft besitzen, daß jede Wurzel einer irreduziblen auflösbaren Gleichung vorgegebenen Grades in ihnen enthalten ist und sie umgekehrt nur solche Werte annehmen, die irreduziblen auflösbaren Gleichungen des vorgegebenen Grades genügen. Eine Lösung des Problems in voller Allgemeinheit hat für einen *Primzahlgrad* Kronecker (*Monatsb. d. Berl. Akad.* (1853), 365) ohne Beweis gegeben; der erste völlig durchgeführte Beweis stammt von

H. Weber. Vgl. das zusammenfassende Resultat bei Weber, *Algebra* **1**, 694. Vgl. ferner Wiman, *Acta math.* **27**, 163 (1903). Die gewöhnliche kurze Angabe, daß die Wurzeln jeder im Körper P irreduziblen auflösbaren Gleichung p^{ten} Grades von der Form $A + \sqrt[p]{R_1} + \sqrt[p]{R_2} + \cdots + \sqrt[p]{R_k}$ sind, wobei A dem Körper P angehört und $R_1, R_2, \ldots, R_k$ die Wurzeln einer im Körper P einfachen Abelschen Gleichung k^{ten} Grades sind, wobei k gleich $p-1$ oder $\frac{p-1}{\delta}$, und δ ein Teiler von $p-1$ ist, hat den Nachteil, daß der angegebene Ausdruck p^k- statt p-wertig ist. Die Lösung des Abelschen Problems der Wurzeldarstellung für auflösbare Gleichungen, die nicht vom Primzahlgrad sind, gibt für den Grad 8: Weber, *Algebra* **2**, 383, für den Grad 9, später allgemein für den Grad p^2: Wiman, *Arkiv för mat., Svenska Vetenskapsakademien* **1**, 665 (1904), ebenda **3**, Nr. 27 (1907), *Verhandl. d. 3. Math.-Kongr.*, S. 190 (1905). Zu den auflösbaren Gleichungen neunten Grades gehören die *Gleichungen für die neun Wendepunkte der Kurven 3. Ordnung.* (Vgl. S. 247.) Die Herstellung eines Wurzelausdruckes für die auflösbaren Gleichungen von beliebigem Primzahlpotenzgrad ist eine noch nicht abgeschlossene Aufgabe. Vgl. Bucht, *Arkiv för mat.*, **5**, Nr. 23 (1909).

Für die *Geometrie* wichtig ist die Frage: *Unter welchen Bedingungen können die Wurzeln einer algebraischen Gleichung nur durch Ausziehen von Quadratwurzeln, also bloß mittelst einer Kette quadratischer Gleichungen, gefunden werden. Hierzu ist notwendig und hinreichend, daß die Ordnung der Galoisschen Gruppe der Gleichung eine Potenz von 2 ist. Dann und nur dann, wenn sich die Lösung einer geometrischen Aufgabe auf eine solche Gleichung zurückführen läßt, können die in der Aufgabe gesuchten Größen mittelst Zirkels und Lineals allein gefunden werden* (Descartes, *Géométrie* (1637), livre premier). Von den berühmten geometrischen Problemen, die das Altertum schon beschäftigten, sind das *delische Problem* (Verdoppelung des Würfels) und die *Trisektion eines beliebigen Winkels* mit Hilfe von Zirkel und Lineal allein unlösbar; diese Aufgaben führen nämlich auf irreduzible Gleichungen dritten Grades. Literatur: Die Bücher von Klein und Enriques (vgl. S. 316). Über *die Teilung des Kreises in n gleiche Teile* mit Hilfe von Zirkel und Lineal vgl. S. 315.

§ 14. Gleichungen fünften und höheren Grades. Einige funktionentheoretische und geometrische, nicht algebraisch auflösbare Gleichungen höheren Grades.

Zum Mittelpunkt für die Behandlung der *Gleichungen fünften Grades* kann die *Ikosaedergleichung:* $\frac{H_{20}^3}{1728 f_{12}^5} = Z$ gewählt werden, die ausführlich geschrieben:

$$(-\vartheta^{20} - 1 + 228\vartheta^{15} - 228\vartheta^5 - 494\vartheta^{10})^3 - 1728\vartheta^5 Z(\vartheta^{10} + 11\vartheta^5 - 1)^5 = 0$$

lautet. (Vgl. oben S. 240.) In dem Körper, der außer den rationalen Zahlen noch den Parameter Z enthält, hat die Ikosaedergleichung eine Galoissche Gruppe der Ordnung $4 \cdot 60$. *Nach Adjunktion der fünften Einheitswurzeln wird die Ikosaedergleichung eine Normalgleichung, deren Wurzeln lineare Funktionen einer einzigen sind und durch die bekannten 60 Kollineationen der Ikosaedergruppe $\mathfrak{J}_{60}$ aus ihr hervorgehen. Die Galoissche Gruppe wird die einfache transitive Permutationsgruppe des Grades und der Ordnung 60, die mit der Ikosaedergruppe $\mathfrak{J}_{60}$ isomorph ist.*

Da die Ikosaedergruppe $\mathfrak{J}_{60}$ Tetraederuntergruppen $\mathfrak{T}_{12}$, also Untergruppen des Index 5, besitzt, so existieren für die Ikosaedergleichung rationale Resolventen fünften Grades.

Setzt man $u = \frac{12 f_{12}^2 \cdot f_6}{T}$, $v = \frac{12 f_{12} \cdot H_8}{H_{20}}$, wobei H_8 und f_6 Invarianten der Oktaedergruppe (vgl. S. 239) und f_{12}, T und H_{20} Invarianten der Ikosaedergruppe (vgl. S. 240) sind, so hängen u und v von dem Verhältnis $\vartheta = \frac{x_1}{x_2}$ ab, und die Funktion $Y = mv + nuv$ genügt, wenn m und n willkürliche Größen sind, einer rationalen Resolvente fünften Grades der Ikosaedergleichung, der von Klein (*Ikosaeder*, S. 106) sog. *Hauptresolvente*

$$Y^5 + \frac{5}{Z}\left(8m^3 + 12m^2 n + \frac{6mn^2 + n^3}{1 - Z}\right) Y^2 + \frac{15}{Z}\left(-4m^4 + \frac{6m^2 n^2 + 4mn^3}{1 - Z} + \frac{3n^4}{4(1 - Z)^2}\right) Y + \frac{3}{Z}\left(48m^5 - \frac{40m^3 n^2}{1 - Z} + \frac{15mn^4 + 4n^5}{(1 - Z)^2}\right) = 0.$$

Diese Hauptresolvente hat in dem durch die Gleichungskoeffizienten bestimmten Körper, nachdem man ihm die fünften Einheitswurzeln adjungiert hat — es genügt sogar die Adjunktion von $\sqrt{5}$ —,

die alternierende Permutationsgruppe von fünf Symbolen, also eine mit der $\mathfrak{S}_{60}$ isomorphe Gruppe, zur Galoisschen Gruppe.[1])

Ist eine beliebige Hauptgleichung fünften Grades, d. h. eine Gleichung (vgl. S. 278) *der Form* $Y^5 + 5\alpha Y^2 + 5\beta Y + \gamma = 0$ *gegeben, so sind zur Bestimmung ihrer Wurzeln folgende Schritte erforderlich:*

(1.) Das Ausziehen der Quadratwurzel aus der Diskriminante der Gleichung. Adjungiert man diese Quadratwurzel dem durch die Gleichungskoeffizienten bestimmten Körper, so hat die vorgelegte Hauptgleichung die alternierende Permutationsgruppe von fünf Symbolen zur Galoisschen Gruppe.

(2.) Nach Adjunktion der Quadratwurzel aus der Diskriminante kann man die vorgelegte Hauptgleichung direkt mit der Hauptresolvente identifizieren. Setzt man:

$$\alpha = \frac{8m^3}{Z} + \frac{12m^2n}{Z} + \frac{6mn^2 + n^3}{Z(1-Z)},$$

$$\beta = 3\left(-\frac{4m^4}{Z} + \frac{6m^2n^2 + 4mn^3}{Z(1-Z)} + \frac{3n^4}{4Z(1-Z)^2}\right),$$

$$\gamma = 3\left(\frac{48m^5}{Z} - \frac{40m^3n^2}{Z(1-Z)} + \frac{15mn^4 + 4n^5}{Z(1-Z)^2}\right),$$

so drücken sich m, n, Z aus diesen Gleichungen rational durch α, β, γ und die Quadratwurzel der Diskriminante aus; die numerischen Koeffizienten sind abgesehen von $\sqrt{5}$ rationale Zahlen (vgl. Klein, *Ikosaeder*, S. 192).

(3.) Bestimmung einer Wurzel ϑ der Ikosaedergleichung $\frac{H_{20}^3}{1728 f_{12}^5} = Z$, deren rechte Seite nach (2) zu finden ist.

(4.) $Y = mv(\vartheta) + nu(\vartheta)v(\vartheta)$ liefert die Wurzeln der vorgelegten Hauptgleichung, hierbei bedeuten m und n die nach (2) bestimmten Werte und u und v sind die oben angegebenen Funktionen von ϑ. *Um die Wurzeln einer Hauptgleichung fünften Grades zu bestimmen, sind also außer rationalen Operationen und dem Ziehen der Quadratwurzel aus der Zahl* 5 (*Operation* (2)) *nur das Ausziehen der Quadratwurzel aus der Diskriminante* (*Operation* (1)) *und die Bestimmung einer Wurzel der Ikosaedergleichung* (*Operation* (3)) *erforderlich.*

---

1) Die Hauptresolvente ist auch ohne Adjunktion von $\sqrt{5}$ eine rationale Resolvente der Ikosaedergleichung; bei Unterlassung der Adjunktion von $\sqrt{5}$ hat sie die symmetrische Gruppe von fünf Buchstaben zur Galoisschen Gruppe.

Um die Operation (3) zu leisten, muß man den algebraischen Bereich verlassen; denn die *Bestimmung der Ikosaederirrationalität ϑ ist transzendenter Natur. Sie kann mittels hypergeometrischer Reihen oder elliptischer Modulfunktionen geschehen.*

Die *hypergeometrische Reihe*

$$x = F(\alpha, \beta, \gamma, Z) = 1 + \frac{\alpha \cdot \beta}{1 \cdot \gamma} Z + \frac{\alpha \cdot (\alpha+1) \cdot \beta \cdot (\beta+1)}{1 \cdot 2 \cdot \gamma \cdot (\gamma+1)} Z^2 + \cdots$$

genügt der Gaußschen Differentialgleichung:

$$Z(1-Z)\frac{d^2x}{dZ^2} + (\gamma - (\alpha+\beta+1)Z)\frac{dx}{dZ} - \alpha\beta x = 0.$$

Der Quotient $\vartheta = \frac{x_1}{x_2}$ irgend zweier partikulärer Integrale dieser linearen homogenen Differentialgleichung befriedigt die Differentialgleichung 3. Ordnung:

$$\frac{\vartheta'''}{\vartheta'} - \frac{3}{2}\left(\frac{\vartheta''}{\vartheta'}\right)^2 = \frac{1-\frac{1}{\nu_1^2}}{2(1-Z)^2} + \frac{1-\frac{1}{\nu_2^2}}{2Z^2} - \frac{\frac{1}{\nu_1^2} - \frac{1}{\nu_3^2} + \frac{1}{\nu_2^2} - 1}{2Z(1-Z)};$$

hierbei bedeuten, falls α, β, γ reell vorausgesetzt werden, die Größen $\frac{1}{\nu_1}, \frac{1}{\nu_2}, \frac{1}{\nu_3}$ der Reihe nach die absoluten Beträge von $|\gamma - \alpha - \beta|$, $|1-\gamma|$ und $|\alpha - \beta|$. Das allgemeine Integral der Differentialgleichung lautet $\frac{a\vartheta + b}{c\vartheta + d}$, wobei a, b, c, d Integrationskonstanten sind. Jedes partikuläre Integral ϑ der Differentialgleichung 3. Ordnung vermittelt die konforme Abbildung der positiven Halbebene der Variablen Z auf die Fläche ϑ eines von Kreisen begrenzten Dreiecks mit den Winkeln $\frac{\pi}{\nu_1}, \frac{\pi}{\nu_2}, \frac{\pi}{\nu_3}$. Das spezielle Wertsystem $\nu_1 = 2$, $\nu_2 = 3$, $\nu_3 = 5$, bei dem die Gaußsche Differentialgleichung der hypergeometrischen Funktion

$$F\left(\frac{11}{60}, -\frac{1}{60}, \frac{2}{3}, Z\right)$$

entspricht, steht in innigster Beziehung zu der durch ein Ikosaeder bewirkten Kugeleinteilung in 120 abwechselnd kongruente und symmetrische sphärische Dreiecke mit den Winkeln $\frac{\pi}{2}, \frac{\pi}{3}, \frac{\pi}{5}$. Die Funktion $Z = \frac{H_{20}^3}{1728 f_{12}^5}$ vermittelt nämlich die konforme Abbildung der Fläche ϑ eines sphärischen Dreiecks mit den Winkeln $\frac{\pi}{2}, \frac{\pi}{3}, \frac{\pi}{5}$ auf die positive Halbebene der Z. *Die Ikosaederirrationalität läßt sich daher als Quotient zweier geeigneter partikulärer Lösungen der Gaußschen Differentialgleichung finden, also mittels hyper-*

geometrischer Reihen bestimmen. (Vgl. die grundlegende Arbeit von H. A. Schwarz, *Journ. f. Math.* **75**, 292 (1873), *Ges. Abhdlgn.*, Berlin 1890, **2**, 211, Klein, *Math. Ann.* **12**, 512 (1877).)

Um die *Ikosaederirrationalität durch elliptische Modulfunktionen* auszudrücken, setze man $Z = J(\omega) = \frac{g_2^3(\omega_1, \omega_2)}{\Delta(\omega_1, \omega_2)}$, wobei J die absolute Invariante, g_2 und Δ homogene transzendente Funktionen $(-4)^{\text{ter}}$ und $(-12)^{\text{ter}}$ Dimension aus der Theorie der elliptischen Funktionen sind, und bestimme hieraus ω. (Wie dies geschehen kann, vgl. Klein, *Math. Ann.* **14**, 111 (1879).) Ist $q = e^{i\pi\omega} = e^{-\pi\frac{K'}{K}}$ in Jacobischer Bezeichnung, so ergibt sich die Ikosaederirrationalität ϑ als Quotient zweier elliptischer Thetareihen in der Form:

$$\vartheta = -\sqrt[5]{q}\,\frac{\vartheta_1(2\omega\pi, q^5)}{\vartheta_1(\pi\omega, q^5)}$$

(Klein, *Math. Ann.* **61**, 560 (1905)). Die Benützung dieser Formel an Stelle von Reihen, die der Gaußschen Differentialgleichung genügen, kommt auf einen Umweg hinaus, wie wenn man sich zur Lösung der reinen Gleichung $z^n = a$ des Logarithmus bedient, zuerst $\frac{1}{n} la$ berechnet und dann z als Numerus zu $\frac{1}{n} la$ bestimmt (vgl. S. 254).

Da die Ikosaedergruppe Diederuntergruppen $\mathfrak{D}_{10}$, also Untergruppen des Index 6, besitzt, so existieren für die Ikosaedergleichung auch rationale Resolventen sechsten Grades. Setzt man $\varphi = \frac{5\,x_1^2 x_2^2 H_{20}}{12 f_{12}^2}$, so genügt die von dem Quotienten $\vartheta = \frac{x_1}{x_2}$ abhängige Größe φ der Gleichung (vgl. Klein, *Ikosaeder*, S. 134f.):

$$\varphi^6 - 10Z\varphi^3 + 12Z^2\varphi + 5Z^2 = 0.$$

Die Galoissche Gruppe dieser Gleichung sechsten Grades ist bei Adjunktion der fünften Einheitswurzeln — es genügt sogar schon die Adjunktion von $\sqrt{5}$ — die der Primzahl $p = 5$ entsprechende Modulargruppe $(p+1)^{\text{ten}}$ Grades der Ordnung $\frac{p(p^2-1)}{2}$, welche mit $\mathfrak{Z}_{60}$ holoedrisch isomorph ist (vgl. S. 253 u. 214).[1]) Setzt man $Z = \frac{g_2^3}{\Delta}$ und $\xi^2 = -\frac{\varphi}{g_2}$, so erhält man für

1) Ohne Adjunktion von $\sqrt{5}$ ist die Gleichung sechsten Grades für φ auch eine rationale Resolvente der Ikosaedergleichung; ihre Galoissche Gruppe ist bei Unterlassung der Adjunktion von der Ordnung 120.

die Unbekannte ξ die in der Theorie der elliptischen Funktionen auftretende Gleichung:

$$\xi^{12} + \frac{10}{\Delta}\xi^6 - \frac{12 g_2}{\Delta^2}\xi^2 + \frac{5}{\Delta^2} = 0,$$

die durch $\frac{1}{\wp\left(\frac{2\omega}{5}\right) - \wp\left(\frac{4\omega}{5}\right)}$ befriedigt wird. (Kiepert, *Journ. f. Math.* 87, 114 (1879).) Führt man in dieser Gleichung $\xi^2 = u$ ein, so erhält man:

$$u^6 + \frac{10}{\Delta}u^3 - \frac{12 g_2}{\Delta^2}u + \frac{5}{\Delta^2} = 0.$$

Diese Gleichung gehört zum Typus der sog. *allgemeinen Jacobischen Gleichungen sechsten Grades. Hierunter versteht man Gleichungen sechsten Grades, bei denen sich die Quadratwurzeln ihrer sechs mit* $u_\infty, u_0, u_1, u_2, u_3, u_4$ *bezeichneten Wurzeln als lineare Funktionen von drei Parametern* A_0, A_1, A_2 *in der bereits von Jacobi* (*Journ. f. Math.* 3, 308 (1828), *Ges. Werke* 1, 261) *angegebenen Form:*

$$\sqrt{u_\infty} = \sqrt{5}A_0, \quad \sqrt{u_m} = A_0 + \varepsilon^m A_1 + \varepsilon^{4m} A_2 \qquad (m = 0, 1, 2, 3, 4)$$

darstellen lassen, wobei ε *eine primitive fünfte Einheitswurzel ist. Die von Brioschi* (vgl. die Angaben auf S. 255, sowie die Darstellung bei Klein, *Ikosueder*, S. 150) *aufgestellte allgemeine Jacobische Gleichung sechsten Grades mit den Wurzeln* $u_\infty, u_0, u_1, u_2, u_3, u_4$ *lautet:*

$$(u - A)^6 - 4A(u - A)^5 + 10B(u - A)^3 - C(u - A) + 5B^2 - AC = 0;$$

hierbei ist:

$$A = A_0^2 + A_1 A_2,$$

$$B = 8A_0^4 A_1 A_2 - 2A_0^2 A_1^2 A_2^2 + A_1^3 A_2^3 - A_0(A_1^5 + A_2^5),$$

$$C = 320A_0^6 A_1^2 A_2^2 - 160A_0^4 A_1^3 A_2^3 + 20A_0^2 A_1^4 A_2^4 + 6A_1^5 A_2^5 - 4A_0(32A_0^4 - 20A_0^2 A_1 A_2 + 5A_1^2 A_2^2)\cdot(A_1^5 + A_2^5) + A_1^{10} + A_2^{10}.$$

Eliminiert man die Parameter A_0, A_1, A_2 aus der oben für die Gleichungswurzeln gegebenen Darstellung, so erhält man folgende zwischen den Wurzeln bestehende Relationen:

$$\sqrt{u_0}+\sqrt{u_1}+\sqrt{u_2}+\sqrt{u_3}+\sqrt{u_4}=\sqrt{5u_\infty},$$
$$\sqrt{u_0}+\varepsilon^2\sqrt{u_1}+\varepsilon^4\sqrt{u_2}+\varepsilon\sqrt{u_3}+\varepsilon^3\sqrt{u_4}=0,$$
$$\sqrt{u_0}+\varepsilon^3\sqrt{u_1}+\varepsilon\sqrt{u_2}+\varepsilon^4\sqrt{u_3}+\varepsilon^2\sqrt{u_4}=0.$$

Der Nachweis, wie man aus der allgemeinen Gleichung fünften Grades nach Adjunktion der Quadratwurzel ihrer Diskriminante allgemeine Jacobische Gleichungen sechsten Grades als rationale Resolventen herleiten kann, und der Zusammenhang spezieller Jacobischer Gleichungen mit der Theorie der elliptischen Funktionen bildet den Kern der Kroneckerschen und Brioschischen Untersuchungen über die Gleichungen fünften Grades. Eine Jacobische Gleichung mit $A=0$ hat Kronecker seiner Auflösung der Gleichung fünften Grades (*C. R.* (1858)) zugrunde gelegt. (Vgl. Klein, *Math. Ann.* **14**, 147 (1879).) $A=0$, $B=\frac{1}{\Delta}$, $C=+\frac{12g_2}{\Delta^2}$ liefert auch die oben angeführte spezielle Gleichung sechsten Grades aus der Theorie der elliptischen Funktionen, von der wir ausgingen. Die Jacobischen Gleichungen für $B=0$ charakterisieren nach Brioschi die aus der Theorie der elliptischen Funktionen stammenden *Multiplikatorgleichungen*. Die spezielle $B=0$, $A=1$, $C=-256k^2(1-k^2)$ entsprechende Jacobische Gleichung verwendet Brioschi zur Lösung der Gleichung fünften Grades (vgl. S. 255).

Die allgemeine Jacobische Gleichung sechsten Grades hängt von zwei wesentlichen Parametern, nämlich $\frac{B}{A^3}$ und $\frac{C}{A^5}$ ab, wie ihre Form:

$$\left(\frac{u}{A}-1\right)^6-4\left(\frac{u}{A}-1\right)^5+10\frac{B}{A^3}\left(\frac{u}{A}-1\right)^3-\frac{C}{A^5}\left(\frac{u}{A}-1\right)+5\left(\frac{B}{A^3}\right)^2-\frac{C}{A^5}=0$$

lehrt, wenn man $\frac{u}{A}=U$ setzt. Hingegen hängt die Ikosaedergleichung nur von einem Parameter ab. Das gleiche trifft auch für die speziellen Jacobischen Gleichungen zu, die $A=0$ bezw. $B=0$ entsprechen und die wir durch die elliptischen Funktionen beherrschen. Die $A=0$ entsprechende Jacobische Gleichung sechsten Grades lautet:

$$u^6+10Bu^3-Cu+5B^2=0$$

oder

$$\varphi^6-10Z\varphi^3+12Z^2\varphi+5Z^2=0,$$

wenn $\varphi = -\frac{C}{12B^2}u$, $Z = \frac{C^3}{12^3B^5}$ gesetzt wird. Die letztere Gleichung ist die Ausgangsgleichung auf Seite 329.

Die Hauptgleichung $Y^5 + 5\alpha Y^2 + 5\beta Y + \gamma = 0$ *hat, wenn die fünften Einheitswurzeln als bekannt gelten, nach Adjunktion der Quadratwurzel ihrer Diskriminante die Ikosaedergleichung zur rationalen Resolvente.* Für die *allgemeine* Gleichung fünften Grades gilt folgender von Kronecker (*Monatsber. d. Berl. Akad.* (1861), 609) stammende und zuerst von Klein (*Ikosaeder*), dann von Gordan (*Math. Ann.* **29**, 318 (1887), vgl. auch Weber, *Algebra* **2**, 470) bewiesene Satz: *Ohne Einführung akzessorischer Irrationalitäten ist es, wenn auch alle numerischen Konstanten als bekannt gelten, selbst nach Adjunktion der Quadratwurzel aus der Diskriminante unmöglich, aus der allgemeinen Gleichung fünften Grades Jacobische Gleichungen mit nur einem Parameter oder überhaupt rationale Resolventen mit nur einem Parameter abzuleiten. Um die Wurzeln einer allgemeinen Gleichung fünften Grades zu finden, kann man die vorgelegte Gleichung zuerst durch Ziehen einer Quadratwurzel auf eine Hauptgleichung reduzieren* (vgl. S. 278). *Dieses Ziehen einer Quadratwurzel geht bei einer allgemeinen Gleichung fünften Grades den mit* (1)—(4) *numerierten Operationen voraus, die wir S. 327 zur Bestimmung der Wurzeln einer Hauptgleichung fünften Grades angaben. Diese bei der Behandlung einer allgemeinen Gleichung fünften Grades noch hinzutretende Quadratwurzel ist eine akzessorische Irrationalität* (vgl. S. 301); *diese trägt nicht zur Erniedrigung der Galoisschen Gruppe der Gleichung bei. Eine solche akzessorische Irrationalität ist nach dem Kroneckerschen Satze nicht zu vermeiden, wenn man mit einparametrigen Gleichungen arbeiten will.* Bei der Lösung der allgemeinen Gleichung fünften Grades durch Jacobische Gleichungen, die $A = 0$ bezw. $B = 0$ entsprechen, tritt daher auch notwendigerweise eine akzessorische Irrationalität auf.

Für die Gleichungen fünften Grades vgl. man außer der im voraufgehenden und auf S. 253 ff. zitierten Literatur den Artikel von Wiman, *Endliche Gruppen linearer Substitutionen in der Enzyklopädie der math. Wiss.* **1**, 522, ferner Scheibner, *Beiträge zur Theorie der linearen Transformationen*, Leipzig 1907. Zur Einführung vgl. man: Vivanti, *Funzioni poliedriche e modulari*, Mailand 1906, Netto, *Algebra* **2**, 487, Bianchi, *Lezioni dei gruppi etc.*, p. 234.

Hat man eine *Gleichung beliebigen Grades*, so kann man sie nach der Galoisschen Theorie zunächst stets durch eine Kette

von Gleichungen mit einfachen Galoisschen Gruppen ersetzen (vgl. S. 303). *Für die Behandlung einer Gleichung $f(z) = 0$ beliebigen Grades mit einfacher Galoisscher Gruppe $\mathfrak{G}$ hat* Klein, (*Math. Ann.* **15**, 251 (1879), vgl. auch Weber, *Algebra* **2**, 235) *folgendes Verfahren angegeben:* Für die Gruppe $\mathfrak{G}$ gibt es, wie für jede Gruppe eine Gruppe linearer homogener Substitutionen von niedrigstem Grade μ, die mit $\mathfrak{G}$ isomorph ist. (*Kleinsches Normalproblem,* vgl. S. 232.) Aus den Wurzeln der vorgelegten Gleichung $f(z) = 0$ kann man stets ganze homogene Funktionen bilden, und zwar genau in der oben definierten Zahl μ, daß sich diese, wenn die Wurzeln von $f(z) = 0$ den Permutationen der Gruppe $\mathfrak{G}$ unterworfen werden, nach einer mit $\mathfrak{G}$ isomorphen linearen homogenen Substitutionsgruppe Γ des Grades μ transformieren. (Vgl. auch Burkhardt, *Math. Ann.* **41**, 309 (1893).) Mit der Gruppe Γ ist wie mit jeder Gruppe linearer homogener Substitutionen das auf S. 228 besprochene sog. *Kleinsche Formenproblem* verbunden, nämlich aus den Invarianten der Gruppe Γ die Variablen zu berechnen. Hierbei gelangt man zu einer Normalgleichung, deren Grad gleich der Ordnung von Γ ist und deren Galoissche Gruppe mit Γ holoedrisch isomorph ist. Die Behandlung der Gleichung $f(z) = 0$ wird durch die angegebene Normalgleichung oder eine andere mit ihr äquivalente rationale Resolvente ersetzt. Existiert eine zu der Gruppe $\mathfrak{G}$ holoedrisch isomorphe Kollineationsgruppe des Grades μ', wobei μ' kleiner als μ ist, und will man die Behandlung der Gleichung $f(z) = 0$ mittels dieser Kollineationsgruppe durchführen, so wird die Adjunktion akzessorischer Irrationalitäten erforderlich.

Hat man eine *einfache Abelsche Gleichung* $f(z) = 0$ (vgl. S. 306) mit den Wurzeln $\alpha_1, \alpha_2, \ldots, \alpha_n$, deren Galoissche Gruppe $\mathfrak{G}$ aus der zyklischen Permutation $C = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ und ihren Potenzen besteht, so wird die fragliche lineare homogene Substitutionsgruppe Γ vom Grade $\mu = 1$. Ist ε eine primitive n^{te} Einheitswurzel und adjungiert man diese dem durch die Koeffizienten von $f(z) = 0$ bestimmten Körper, so stellt die in dem erweiterten Körper natürliche Irrationalität $Y_1 = (\alpha_1 + \varepsilon\alpha_2 + \varepsilon^2\alpha_3 + \cdots + \varepsilon^{n-1}\alpha_n)^{n-1}$ eine ganze rationale Funktion der Wurzeln von $f(z) = 0$ dar, die bei der Permutation $C^\varkappa$ in $\varepsilon^\varkappa Y_1$ $(\varkappa = 1, 2, \ldots, n)$ übergeht. Einfacher verwendet man die sich mit Y_1 kogredient transformierende Größe $Z_1 = \dfrac{1}{\alpha_1 + \varepsilon\alpha_2 + \varepsilon^2\alpha_3 + \cdots + \varepsilon^{n-1}\alpha_n}$. Die Invariante dieser Gruppe in einer Variablen ist $Z_1{}^n$. Die Bestimmung von Z_1 kommt auf das Ziehen einer n^{ten} Wurzel aus

der bekannten Größe 1: $(\alpha_1 + \varepsilon\alpha_2 + \varepsilon^2\alpha_3 + \cdots + \varepsilon^{n-1}\alpha_n)^n$ heraus. Die Größe $\frac{1}{Z_1}$ ist die Lagrangesche Resolvente (vgl. S. 307). *Die durch Wurzelzeichen lösbaren Gleichungen und auch nur sie allein führen auf Formenprobleme ersten Grades.*

Bei der nicht algebraisch auflösbaren Gleichung niedrigsten Grades, der *Gleichung fünften Grades*, liegt es so: Nach Adjunktion der Quadratwurzel aus der Diskriminante hat sie die alteenierende Gruppe von fünf Symbolen zur Galoisschen Gruppe. Dr. Gruppe linearer homogener Substitutionen niedrigsten Gradeis die mit dieser einfachen Gruppe der Ordnung 60 isomorph islt ist eine solche des Grades 3, die ternäre Ikosaedergruppe (vgl. S. 242). Das Kleinsche Formenproblem wird ungefähr gleichbedeutend mit dem Studium der allgemeinen Jacobischen Gleichungen sechsten Grades. (Klein, *Math. Ann.* **15**, 259 (1879), *Ikosaeder*, S. 211). Da es keine einfache binäre Gruppe linearer homogener Substitutionen 60. Ordnung gibt, erfordert die Behandlung der Gleichung fünften Grades mit Hilfe der binären Kollineationsgruppe $\mathfrak{J}_{60}$ die Zuziehung einer akzessorischen Quadratwurzel.

Die *allgemeine Gleichung sechsten Grades* hat nach Adjunktion der Quadratwurzel aus der Diskriminante die alternierende Gruppe von sechs Symbolen zur Galoisschen Gruppe. Die Valentinersche Gruppe (vgl. S. 242) ist die Kollineationsgruppe niedrigsten Grades 3, die mit der einfachen Gruppe der Ordnung 360 isomorph ist. Die Valentinersche Gruppe ist mit keiner Gruppe linearer homogener Substitutionen in drei Variablen isomorph. Unter Zuziehung akzessorischer Irrationalitäten läßt sich die Auflösung der Gleichung sechsten Grades mit der Valentinerschen Gruppe in Zusammenhang bringen (vgl. Klein, *Journ. f. Math.* **129**, 151 (1905), Gordan, *Math. Ann.* **61**, 453 (1905), Lachtin, *Math. Ann.* **56**, 445 (1903)).

Ehe die Valentinersche Gruppe entdeckt war, haben Maschke (*Rend. Lincei* 4_1, 181 (1888)) und Brioschi, *Acta math.* **12**, 83 (1889) und *Ann. de l'école normale* (3) **12**, 343 (1895) die Wurzeln der Gleichung sechsten Grades durch hyperelliptische ϑ-Nullwerte ausgedrückt, analog wie man die Wurzeln der Gleichung fünften Grades durch elliptische Modulfunktionen bestimmen kann. Der tiefere Grund hierfür liegt darin, daß die Gruppe der Borchardtschen Moduln $\mathfrak{G}_{11520}$ eine invariante Untergruppe der Ordnung 16 besitzt und daher mit der symmetrischen Gruppe $\mathfrak{S}_{6!}$ von sechs Symbolen meroedrisch isomorph ist. (Vgl. S. 243.)

Da die alternierende Gruppe $\mathfrak{A}_{\frac{7!}{2}}$ von sieben Symbolen mit einer quaternären Kollineationsgruppe holoedrisch isomorph ist, so kann die *allgemeine Gleichung siebenten Grades* mit einer quaternären Kollineationsgruppe in Zusammenhang gebracht werden (vgl. S. 236 u. 243, sowie Klein, *Math. Ann.* 28, 499 (1887)). Auch die *Gleichung sechsten Grades* kann mit einer quaternären Kollineationsgruppe behandelt werden, da die symmetrische Gruppe $\mathfrak{S}_{6!}$ von sechs Symbolen mit einer solchen holoedrisch isomorph ist (vgl. Klein, *Math. Ann.* 28, 499 (1887).) Für die Gleichung siebenten Grades ist die quaternäre Kollineationsgruppe die niedrigste mögliche, hingegen nicht für die Gleichung sechsten Grades (vgl. oben). Sowohl die mit $\mathfrak{S}_{6!}$ als auch mit $\mathfrak{A}_{\frac{7!}{2}}$ holoedrisch isomorphen quaternären Kollineationsgruppen erfordern, wenn man zu einer homogenen linearen Substitutionsgruppe übergeht, die doppelte Zahl von Substitutionen. Die Behandlung der Gleichungen sechsten und siebenten Grades auf diesem Wege erfordert daher akzessorische Irrationalitäten.

Die allgemeinen Gleichungen achten und höheren Grades können, wenn n der Grad der Gleichung ist, mit keiner Kollineationsgruppe von niedrigerem als $(n-1)^{ten}$ *Grade in Zusammenhang gebracht werden* (Folge des Satzes von Wimann, vgl. S. 235). Die allgemeine Gleichung n^{ten} Grades läßt sich sofort auf ein Formenproblem $(n-1)^{ten}$ Grades reduzieren, indem man die zu permutierenden Größen der Bedingung unterwirft, daß ihre Summe verschwinden soll. *Die allgemeine Gleichung* n^{ten} *Grades gestattet daher für* $n \geqq 8$ *keine Reduktion auf einfachere Formenprobleme; die fraglichen Gleichungen bilden ihre eigenen Normalprobleme.*

Eine Bestimmung der Wurzeln der allgemeinen Gleichung n^{ten} Grades durch transzendente Funktionen geben die Aufsätze von Lindemann, *Gött. Nachr.* (1884) und (1892).

Über *spezielle, nicht algebraisch auflösbare Gleichungen* mögen folgende Angaben gemacht werden: *Daß gewisse aus der Theorie der elliptischen Funktionen stammende Gleichungen* $(p+1)^{ten}$ *Grades* (p *ungerade Primzahl*) *die Modulargruppe der Ordnung* $\frac{1}{2}p(p^2-1)$ *als Galoissche Gruppe* besitzen, hat bereits Galois (*Œuvres*, p. 27) gezeigt (vgl. S. 253). Kronecker (*Monatsber. d. Berl. Akad.* (1861) 615, (1879) 220) hat dann die Aufmerksamkeit auf die *allgemeinen Gleichungen* $(p+1)^{ten}$ *Grades* (p *ungerade Primzahl*) *mit der Modulargruppe als Galoisscher Gruppe* gelenkt. Für $p = 3, 5, 7$ und 11 und auch nur für diese Werte

haben die fraglichen Gleichungen vierten, sechsten, achten und zwölften Grades infolge der Eigenschaften der Modulargruppe (vgl. S. 215, letzte Zeile) rationale Resolventen dritten, fünften, siebenten und elften Grades. Die Behandlung jeder Gleichung siebenten und achten Grades, deren Galoissche Gruppe eine einfache Gruppe der Ordnung 168 ist ($p = 7$ entsprechend), läßt sich auf das Formenproblem einer ternären linearen homogenen Substitutionsgruppe der Ordnung 168 zurückführen (vgl. S. 242). Ebenso kann die Behandlung jeder Gleichung 11. und 12. Grades mit einer einfachen Galoisschen Gruppe der Ordnung 660 ($p = 11$ entsprechend) auf das Formenproblem einer quinären linearen homogenen Substitutionsgruppe zurückgeführt werden; die Kollineationsgruppe niedrigsten Grades, die mit der $p = 11$ entsprechenden Modulargruppe isomorph ist, geht nämlich aus einer quinären Gruppe linearer homogener Substitutionen derselben Ordnung hervor (vgl. S. 244). Für alle Gleichungen siebenten und achten Grades mit einer einfachen Galoisschen Gruppe der Ordnung 168 hat Klein (*Math. Ann.* **15**, 265 (1879)) bewiesen, daß sie sich unter Heranziehung einer akzessorischen Hilfsgleichung vierten Grades mittels der speziellen, durch die Theorie der elliptischen Funktionen gelieferten Modulargleichung lösen lassen. (Rechnerische Ausführung bei Gordan, Literatur auf S. 242 bei der Kleinschen Gruppe.)

Die Dreiteilung der hyperelliptischen Funktionen liefert eine Gleichung 27. Grades mit einer Galoisschen Gruppe der Ordnung 51840, die eine invariante einfache Untergruppe der Ordnung 25920 besitzt. Die Lösung jeder Gleichung 27. Grades mit einer derartigen Galoisschen Gruppe der Ordnung 51840 läßt sich auf die obige Gleichung aus der Theorie der hyperelliptischen Funktionen zurückzuführen. Zu dieser Gattung von Gleichungen gehört im besonderen auch die *Gleichung 27. Grades für die 27 Geraden einer allgemeinen Fläche dritter Ordnung.* (Literaturangaben oben auf S. 244.)

Die Gleichung 16. Grades für die 16 Knotenpunkte einer Kummerschen Fläche hat eine Galoissche Gruppe der Ordnung 11520, die eine invariante Untergruppe $\mathfrak{G}_{16}$ der Ordnung 16 besitzt. Die Quotientengruppe $\mathfrak{G}_{11520}/\mathfrak{G}_{16}$ ist mit der symmetrischen Permutationsgruppe in sechs Symbolen holoedrisch isomorph, und die Gleichung 16. Grades kann mit Hilfe der allgemeinen Gleichung sechsten Grades und vier Quadratwurzelausziehungen gelöst werden (Literatur oben auf S. 243). Hingegen erfordert die *Gleichung 16. Grades für die 16 Geraden einer Fläche*

4. Ordnung mit einem Doppelkegelschnitt außer der Auflösung quadratischer Gleichungen nur die Lösung einer allgemeinen Gleichung fünften Grades. Die Ordnung der Galoisschen Gruppe dieser Gleichung 16. Grades ist 16. 5!. (Jordan, *Traité*, p. 309, Klein, *Math. Ann.* **4**, 357 (1871), Pereno, *Annali di mat.* (2) **21**, 57 (1893).)

Die *28 Doppeltangenten einer allgemeinen Kurve 4. Ordnung* hängen von einer Gleichung 28. Grades mit einer einfachen Galoisschen Gruppe der Ordnung 1451520 ab (vgl. S. 249). Die Verallgemeinerung dieses Problems ist folgende: *Ist C_n eine Kurve n^{ter} Ordnung ohne Doppelpunkte vom Geschlechte $p = \frac{1}{2}(n-1)(n-2)$, so hängt die Bestimmung der $2^{p-1}(2^p-1)$ Kurven der Ordnung $n-3$, die die gegebene C_n in $\frac{1}{2}n(n-3)$ Punkten zweipunktig berühren, von einer Gleichung $2^{p-1}(2^p-1)^{ten}$ Grades ab.* Die Untersuchung der zugehörigen Galoisschen Gruppe bei Jordan, *Traité*, p. 329, Dickson, *Trans. Am. M. S.* **3**, 38 und 377 (1902), *Annals of math.* (2) **6**, 141 (1905). Weiteres über Gruppen geometrischer Gleichungen bei Jordan, *Traité*, p. 301, Maillet, *C. R.* **138**, 890 (1904), *Ann. de Toulouse* (2) **6**, 277 (1904), E. Pascal, *Annali di mat.* (2) **20**, 269 (1893), **21**, 85 (1893).

§ 15. Bestimmung der Anzahl der reellen und komplexen Wurzeln einer Gleichung.

Eine Gleichung $f(z) = a_0 z^n + a_1 z^{n-1} + \cdots + a_n = 0$ mit reellen Koeffizienten hat zwischen zwei reellen Zahlen A und B eine gerade (einschl. 0) oder eine ungerade Anzahl von Wurzeln, je nachdem $f(A)$ und $f(B)$ gleiche oder entgegengesetzte Vorzeichen haben; hierbei ist jede Wurzel entsprechend ihrer Vielfachheit zu zählen. Für reelle Größen A, deren absoluter Betrag genügend groß ist, entscheidet über das Vorzeichen von $f(A)$ das erste Glied $a_0 A^n$.

Eine Gleichung ungeraden Grades mit reellen Koeffizienten hat stets wenigstens eine reelle Wurzel. Ist[1] *$sign(a_0) = -sign(a_n)$, so hat die Gleichung wenigstens eine positive Wurzel und, wenn n eine gerade Zahl ist, auch noch wenigstens eine negative Wurzel.*

Ist $f(z)$ eine ganze Funktion mit reellen Koeffizienten, $f'(z)$ ihre erste Abgeleitete und durchläuft z die reelle Zahlenreihe von

1) sign = Vorzeichen.

$-\infty$ *bis* $+\infty$, *so geht der Quotient* $\frac{f(z)}{f'(z)}$ *bei seinem Verschwinden jedesmal von einem negativen Wert zu einem positiven über.*

Hieraus folgt der Satz von Rolle, *Traité d'algèbre* (1690), 2tes Buch, Kap. 6, vgl. auch Lagrange, *Œuvres* 8, 190—199: *Sind A und B zwei unmittelbar aufeinanderfolgende reelle Wurzeln einer Gleichung* $f(z) = 0$ *mit reellen Koeffizienten, so liegt zwischen A und B (die Grenzen sind nicht mitzuzählen) eine ungerade Zahl von Wurzeln (mehrfache Wurzeln sind ihrer Vielfachheit nach zu zählen) der Gleichung* $f'(z) = 0$, *also mindestens eine.*

Korollare zum Satz von Rolle: I. Hat die Gleichung $f(z) = 0$ mit reellen Koeffizienten ϱ reelle Wurzeln, so hat die Gleichung $f'(z) = 0$ stets $\varrho - 1 + 2k$ reelle Wurzeln, wobei $k \geqq 0$ ist. II. *Zwischen zwei aufeinanderfolgenden reellen Wurzeln der Gleichung* $f'(z) = 0$ *kann höchstens eine Wurzel von* $f(z) = 0$ *gelegen sein.* Den Satz II benützt Rolle, um Grenzen für die reellen Gleichungswurzeln zu finden. Vgl. auch Lagrange, a. a. O.

Aus dem Satze von Rolle ergibt sich das Theorem von Waring (*Meditationes algebraicae* (1770), vgl. Cajori bei Cantor, *Vorl. über Geschichte der Math.* 4, 106): *Sind A und B zwei unmittelbar aufeinanderfolgende reelle Wurzeln einer Gleichung* $f(z) = 0$ *mit reellen Koeffizienten, so liegt zwischen A und B (die Grenzen sind nicht mitzuzählen) eine ungerade Anzahl von Wurzeln (mehrfache Wurzeln sind ihrer Vielfachheit nach zu zählen) der Gleichung* $F(z) = C_0 f(z) + C_1 f'(z) = 0$, *also mindestens eine.* C_0 *und* C_1 *sind beliebige reelle Größen;* $C_1 \neq 0$ (vgl. auch Cesàro, *Nouv. Ann. de math.* (3) 4, 326 (1885)).

Korollare zum Waringschen Satz: I. Hat die Gleichung $f(z) = 0$ mit reellen Koeffizienten ϱ reelle Wurzeln, so hat die Gleichung $F(z) = C_0 f(z) + C_1 f'(z) = 0$, wobei C_0 und C_1 beliebige reelle Größen sind $(C_0 \neq 0)$, $\varrho + 2l$ reelle Wurzeln; hierbei ist $l \geqq 0$. II. Zwischen zwei aufeinanderfolgenden reellen Wurzeln der Gleichung $F(z) = 0$ kann höchstens eine Wurzel von $f(z) = 0$ gelegen sein.

Verallgemeinerung des Korollars I zum Waringschen Satz: Hat die Gleichung $C_0 z^m + C_1 z^{m-1} + C_2 z^{m-2} + \cdots + C_m = 0$ von beliebigem Grade m mit reellen Koeffizienten lauter reelle Wurzeln und die Gleichung $f(z) = 0$ vom Grade $n \geqq m$ mit reellen Koeffizienten ϱ reelle Wurzeln, so hat die Gleichung $C_0 f(z) + C_1 f'(z) + C_2 f''(z) + \cdots + C_m f^{(m)}(z) = 0$, bei der $f'(z)$, $f''(z), \ldots f^{(m)}(z)$ die sukzessiven Abgeleiteten bedeuten, $\varrho + 2l$

reelle Wurzeln ($l \geqq 0$) (Poulain, *Nouv. Ann. de math.* (2) **6**, 23 (1867), Realis, ebenda, p. 417, Fouret, *C. R.* **106**, 1135 u. 1220 (1888)).

Hat die Gleichung $f(z) = 0$ vom n^{ten} Grade mit reellen Koeffizienten ϱ reelle Wurzeln, so hat die Gleichung

$$f(z) + tf'(z) + t^2 f''(z) + \cdots + t^n f^{(n)}(z) = 0,$$

wobei t eine beliebige reelle Größe ist und $f'(z)$, $f''(z)$, ..., $f^{(n)}(z)$ die sukzessiven Abgeleiteten bedeuten, $\varrho - 2k$ reelle Wurzeln, $k \geqq 0$ (Poulain, *Nouv. Ann. de math.* (2) **6**, 22 (1867), Realis, ebenda, p. 416).

Satz von Hermite *und* Biehler (Hermite, *Bull. soc. math.* **7**, 131 (1879), Biehler, *Journ. f. Math.* **87**, 350 (1879), Laguerre, *Œuvres* **1**, 109 und 360, Auric, *C. R.* **137**, 967 (1903), Hurwitz, *Math. Ann.* **46**, 284 (1895)): *Ist* $(z-\alpha_1-\beta_1 i)(z-\alpha_2-\beta_2 i)\ldots(z-\alpha_n-\beta_n i) = \varphi(z) + i\psi(z)$ *und haben die Größen* $\beta_1, \beta_2, \ldots, \beta_n$ *ausnahmslos das gleiche Vorzeichen*[1]*, so hat die Gleichung* $p\varphi(z) + q\psi(z) = 0$, *wobei p und q beliebige reelle Größen bedeuten, nur reelle Wurzeln. Alle reellen Wurzeln von $\varphi(z) = 0$ sind untereinander verschieden und werden durch diejenigen von $\psi(z) = 0$ getrennt, die gleichfalls reell und verschieden sind.*

Die *genaue Anzahl reeller Wurzeln* einer beliebigen Gleichung $f(z) = 0$ mit reellen Koeffizienten zwischen zwei reellen Zahlen A und B $(A < B)$ liefert der *Satz von* Sturm (vgl. S. 256). Wir führen zunächst den Begriff der *Sturmschen Kette* ein: Ist $f(z) = 0$ eine Gleichung mit reellen Wurzeln, die im Intervall A bis B keine mehrfachen Wurzeln hat, so heißt eine Reihe ganzer rationaler Funktionen $f(z)$, $f_1(z)$, $f_2(z)$, ..., $f_r(z)$ mit reellen Koeffizienten eine Sturmsche Kette, wenn sie folgende Bedingungen erfüllt:

1. Die letzte Funktion $f_r(z)$ ändert nie ihr Vorzeichen, wenn z alle reellen Werte des Intervalls von A bis B[2]) durchläuft.

2. Zwei unmittelbar aufeinanderfolgende Funktionen $f_{k-1}(z)$ und $f_k(z)$ verschwinden nie gleichzeitig für denselben reellen Wert $z = b$ des Intervalls von A bis B.

3. Verschwindet eine Funktion $f_k(z)$ $(k = 1, 2, \ldots, r-1)$ für einen reellen Wert $z = b$ des Intervalls von A bis B, so ist $\operatorname{sign} f_{k-1}(b) = -\operatorname{sign} f_{k+1}(b)$.

---

1) i ist die imaginäre Einheit.

2) Die Grenzen A und B sind bei den Bedingungen 1—4 einzuschließen.

4. Wenn $f(z)$ für einen reellen Wert des Intervalls von A bis B verschwindet, so hat für diesen die erste Funktion $f_1(z)$ jedesmal dasselbe Vorzeichen wie die erste Abgeleitete $f'(z)$.

Sturmscher Satz (I): *Bilden für eine Gleichung* $f(z) = 0$ *mit reellen Koeffizienten, die im Intervall von* A *bis* B *keine mehrfachen Wurzeln hat, die Funktionen* $f(z), f_1(z), f_2(z), \ldots, f_r(z)$ *eine Sturmsche Kette und sind* A *und* B *zwei reelle Zahlen* $(A < B)$, *so hat die Gleichung* $f(z) = 0$ *zwischen den Grenzen* A *und* B, *die selbst keine Wurzeln sein sollen, soviel reelle Wurzeln, wie die Reihe*

$$f(A), f_1(A), f_2(A), \ldots, f_r(A)$$

mehr Vorzeichenwechsel besitzt als die Reihe

$$f(B), f_1(B), f_2(B), \ldots, f_r(B).$$

Zwei aufeinanderfolgende reelle Größen, die nicht verschwinden, besitzen einen *Vorzeichenwechsel* oder eine *Vorzeichenfolge*, je nachdem sie verschiedene oder gleiche Vorzeichen haben. In den zwei obigen Reihen können Nullen nur in der Mitte, jedoch nie am Ende (Bedingung 1 für die Sturmsche Kette) und nie am Anfang ($f(A) \neq 0$, $f(B) \neq 0$) auftreten; die Nullen kommen nur isoliert vor (Bedingung 2 für die Sturmsche Kette) und sind beim Abzählen der Vorzeichenwechsel einfach fortzulassen.

Eine besondere Sturmsche Kette liefert das *von Sturm angegebene Divisionsverfahren.* Man wähle für $f_1(z)$ die erste Abgeleitete $f'(z)$. Durch fortgesetzte Division bilde man die Gleichungskette:

$$\begin{aligned} f(z) &= f'(z)\,G_1(z) - f_2(z), \\ f'(z) &= f_2(z)\,G_2(z) - f_3(z), \\ f_2(z) &= f_3(z)\,G_3(z) - f_4(z), \\ &\vdots \\ f_{r-2}(z) &= f_{r-1}(z)\,G_{r-1}(z) - f_r, \end{aligned}$$

in der die sukzessiv auftretenden Reste immer mit entgegengesetzten Vorzeichen verwendet werden. Hat $f(z) = 0$ keine mehrfachen Wurzeln, so sind $f(z)$ und $f'(z)$ relativ prim, und man gelangt schließlich zu einem von Null verschiedenen letzten konstanten Rest: $-f_r$. *Die gefundenen Funktionen* $f(z), f'(z), f_2(z), f_3(z), \ldots, f_r$ *bilden eine Sturmsche Kette.*

Sturmscher Satz (II): *Hat die Gleichung* $f(z) = 0$ *reelle Koeffizienten und keine mehrfachen Wurzeln und setzt man in*

die durch das Sturmsche Divisionsverfahren gefundene Funktionenkette $f(z)$, $f'(z)$, $f_2(z)$, $f_3(z)$, ..., f_r *die reellen Zahlen* A *und* B $(A < B)$, *die selbst keine Wurzeln der Gleichung sein sollen, so gibt die Anzahl der Vorzeichenwechsel, welche die Reihe*

$$f(A),\ f'(A),\ f_2(A),\ \ldots,\ f_{r-1}(A),\ f_r$$

mehr als die Reihe

$$f(B),\ f'(B),\ f_2(B),\ \ldots,\ f_{r-1}(B),\ f_r$$

hat, die genaue Anzahl reeller Wurzeln von $f(z) = 0$, *die zwischen den Grenzen* A *und* B *liegen.*

Zusatz (Sturm, a. a. O., Artikel 18): *Bei einer Gleichung mit mehrfachen Wurzeln verschwindet der Rest* $-f_r$; *der Satz* II *bleibt trotzdem noch richtig, nur ist das verschwindende* f_r *fortzulassen und jede zwischen den Grenzen* A *und* B *gelegene mehrfache Wurzel auch bloß einfach zu zählen.*

Wählt man $A = -\infty$ und $B = +\infty$, so liefert das Sturmsche Theorem die genaue Anzahl aller reellen Wurzeln von $f(z) = 0$.

Eine Gleichung $f(z) = 0$ *vom* n^{ten} *Grade mit reellen Koeffizienten besitzt dann und nur dann lauter reelle verschiedene Wurzeln, wenn die durch das Sturmsche Divisionsverfahren gewonnene Sturmsche Kette* f, f', f_2, ..., f_r *aus* $n+1$ *Funktionen besteht, und die Koeffizienten der höchsten Potenzen in allen Funktionen das nämliche Vorzeichen haben.*

Ebenso wie das Sturmsche Verfahren leistet auch die Theorie der reellen quadratischen Formen (Hermite, *C. R.* **35**, 52 (1852), ebenda **36**, 407 (1853), *Journ. f. Math.* **52**, 39 (1856), *Œuvres* **1**, 281, 284, 397, Sylvester, *Phil. Trans.* **143** (1853), *Coll. math. papers* **1**, 429) *die Bestimmung der genauen Anzahl reeller Wurzeln, welche eine Gleichung* $f(z) = a_0 z^n + a_1 z^{n-1} + \cdots + a_n = 0$ *mit reellen Koeffizienten zwischen zwei gegebenen reellen Zahlen* A *und* B $(A < B)$ *hat.*

Es seien $\alpha_1, \alpha_2, \ldots, \alpha_n$ die n verschiedenen Wurzeln der Gleichung $f(z) = 0$, die keine mehrfachen Wurzeln habe. Man bilde die quadratische Form:

$$h_1 = \sum_{k=1}^{k=n} (\xi - \alpha_k)^q (U_1 + \alpha_k U_2 + \alpha_k^2 U_3 + \cdots \alpha_k^{n-1} U_n)^2$$

der n Variablen $U_1, U_2, \ldots, U_n$; die Größe q bedeute eine positive oder negative ungerade Zahl, ξ einen reellen willkürlichen Parameter. Die Koeffizienten von h_1 lassen sich als

symmetrische Funktionen der Gleichungswurzeln durch die reellen Koeffizienten von $f(z) = 0$ reell ausdrücken. *Transformiert man die reelle quadratische Form h_1 durch eine reelle lineare homogene Substitution von nicht verschwindender Determinante in eine Summe reeller Quadrate* (vgl. S. 121) *und findet man, wenn man $\xi = A$ setzt, N_A, wenn man $\xi = B$ wählt, N_B Quadrate mit negativen Vorzeichen, so ist die genaue Anzahl der zwischen A und B gelegenen Wurzeln von $f(z) = 0$ gleich $N_A - N_B$.*

Man kann statt h_1 auch andere reelle quadratische Funktionen wählen, die dasselbe leisten, z. B. die aus h_1 durch reelle lineare homogene Transformation von nicht verschwindender Determinante: $U_i = a_0 u_i + a_1 u_{i+1} + a_2 u_{i+2} + \cdots + a_{n-i} u_n$ $(i = 1, 2, \ldots, n)$ hervorgehende reelle quadratische Form H_1 der n Variablen $u_1, u_2, \ldots u_n$, nämlich:

$$H_1 = \sum_{k=1}^{k=n} (\xi - \alpha_k)^q \cdot [a_0 u_1 + (a_0 \alpha_k + a_1) u_2 + (a_0 \alpha_k^2 + a_1 \alpha_k + a_2) u_3 + \cdots (a_0 \alpha_k^{n-1} + a_1 \alpha_k^{n-2} + \cdots a_{n-1}) u_n]^2.$$

Besonders bequem gestaltet sich die Berechnung der $q = 1$ entsprechenden quadratischen Form H_1. Sie ist nämlich die *Bézoutiante* (vgl. S. 271) der zwei ganzen Funktionen $f(x)$ und $g(x) = f'(x) \cdot (x - \xi)$ vom n^{ten} Grade, wobei $f'(x)$ die erste Abgeleitete bedeutet. Es ergibt sich folgende Vorschrift: *Man entwickle*

$$\frac{f(x) \cdot f'(y) \cdot (y - \xi) - f(y) f'(x) \cdot (x - \xi)}{y - x} = \sum_{i=0}^{i=n-1} \sum_{k=0}^{k=n-1} d_{ik} x^i y^k$$

nach ganzen positiven Potenzen der Variablen x, y und führe für $x^i y^k$ die Produkte $u_{n-i} u_{n-k}$ ein. Die $q = 1$ entsprechende quadratische Form H_1 lautet alsdann

$$H_1 = \sum_{i=0}^{i=n-1} \sum_{k=0}^{k=n-1} d_{ik} u_{n-i} u_{n-k}.$$

Die Differenz der Anzahl negativer Quadrate, welche die Bézoutiante $\sum_{i=0}^{i=n-1} \sum_{k=0}^{k=n-1} d_{ik} u_{n-i} u_{n-k}$, deren Koeffizienten von dem Parameter ξ abhängen, für $\xi = A$ und $\xi = B$ liefert, ist gleich der Zahl der Gleichungswurzeln von $f(z) = 0$ zwischen den Grenzen A und B (Hermite, *Journ. f. Math.* **52**, 51 (1856), *Œuvres* **1**, 413).

Die für $q = 0$ aus h_1 hervorgehende quadratische Form sei mit g_1 bezeichnet.

$$g_1 = \sum_{k=1}^{k=n} (U_1 + \alpha_k U_2 + \alpha_k^2 U_3 + \cdots + \alpha_k^{n-1} U_n)^2 = \sum_{i=1}^{i=n} \sum_{l=1}^{l=n} s_{i+l-2} U_i U_l,$$

wobei $s_i = \sum_{k=1}^{k=n} \alpha_k^i$ die Summe der i^{ten} Potenzen der Gleichungswurzeln bedeutet. *Die reelle quadratische Form g_1 hat bei jeder reellen linearen homogenen Substitution in eine Summe von Quadraten soviel Quadrate mit negativen Vorzeichen als die Gleichung $f(z) = 0$ Paare imaginärer Wurzeln hat.* Hieraus ergibt sich der Satz von Borchardt (*Journ. de math.* **12**, 58 (1847), *Ges. Werke*, S. 24):

Die Gleichung $f(z) = 0$ mit reellen Koeffizienten und verschiedenen Wurzeln hat soviel Paare imaginärer Wurzeln als die Reihe der Determinanten:

$$\sigma_1 = s_0, \quad \sigma_2 = \begin{vmatrix} s_0 & s_1 \\ s_1 & s_2 \end{vmatrix},$$

$$\sigma_3 = \begin{vmatrix} s_0 & s_1 & s_2 \\ s_1 & s_2 & s_3 \\ s_2 & s_3 & s_4 \end{vmatrix}, \ldots, \quad \sigma_n = \begin{vmatrix} s_0 & s_1 \cdots s_{n-1} \\ s_1 & s_2 \cdots s_n \\ \vdots & \\ s_{n-1} & s_n \cdots s_{2n-2} \end{vmatrix}$$

Vorzeichenwechsel aufweist. Das Borchardtsche Theorem bleibt auch richtig, wenn in der Größenreihe $\sigma_1, \sigma_2, \ldots, \sigma_n$ isolierte Nullen auftreten; diese sind einfach zu streichen. Bei mehrfach unmittelbar aufeinanderfolgenden Nullen kann das Borchardtsche Kriterium versagen, wie beispielsweise die Gleichung $z^4 + 1 = 0$ lehrt.

Eine charakteristische Bedingung dafür, daß eine Gleichung $f(z) = 0$ mit reellen Koeffizienten ohne mehrfache Wurzeln nur reelle Wurzeln hat, besteht darin, daß $\sigma_2 > 0$, $\sigma_3 > 0, \ldots, \sigma_n > 0$ sind.

Die quadratische Form g_1 und das Trägheitsgesetz der reellen quadratischen Formen hat Jacobi bereits vor Veröffentlichung der Hermiteschen und Sylvesterschen Arbeiten zur Bestimmung der reellen Wurzeln einer Gleichung $f(z) = 0$ mit reellen Koeffizienten verwendet (vgl. Borchardt, *Journ. f. Math.* **53**, 281 (1857), *Ges. Werke*, S. 469, siehe oben S. 122).

Wir schließen diese Betrachtungen (vgl. zu dem hier behandelten Gegenstand noch die Aufsätze von Kronecker, *Ges. Werke* **1**, 227, 303, **2**, 37, 113, die *Monographie* von Hattendorff, *Die Sturmschen Funktionen*, 2. Aufl. Hannover 1874, Frobenius, *Journ. f. Math.* **114**, 187 (1895), Hurwitz, *Math. Ann.* **46**, 273 (1895), die Anmerkungen zu Sturms Abhandlung in Ostwalds *Klass. der exakten Wiss.* Nr. 143) mit einem Satz von Sylvester (*Phil. Trans.* (1853), *Coll. math. papers* **1**,547):

Sei $f(z) = 0$ *eine Gleichung* n^{ten} *Grades ohne mehrfache Wurzeln und mit reellen Koeffizienten und* $f_1(z) = nf(z) - zf'(z)$, *wobei* $f'(z)$ *die erste Abgeleitete von* $f(z)$ *ist, so hat die Gleichung* $f(z) = 0$ *soviel Paare imaginärer Wurzeln, wie die Bézoutiante von* $f'(z)$ *und* $f_1(z)$ *bei einer reellen Transformation in eine Summe von Quadraten negative Quadrate aufweist. Die Bézoutiante von* $f'(z)$ *und* $f_1(z)$ *ist eine reelle quadratische Form*

$$\sum_{i=0}^{i=n-2} \sum_{k=0}^{k=n-2} d_{ik} u_{n-i} u_{n-k}$$

von nur $n-1$ *Variablen* $u_2, u_3, \ldots, u_n$ *und wird gefunden, indem man*

$$\frac{f'(x) \cdot f_1(y) - f'(y) f_1(x)}{y - x} = \sum_{i=0}^{i=n-2} \sum_{k=0}^{k=n-2} d_{ik} x^i y^k$$

bildet und für $x^i y^k$ *das Produkt* $u_{n-i} u_{n-k}$ *setzt.*

Nicht die genaue Anzahl, aber eine *obere Grenze* für die Zahl der Wurzeln einer Gleichung mit reellen Koeffizienten, die zwischen zwei reellen Zahlen liegen, liefert der *Satz von* Fourier (vgl. S. 256): *Ist* $f(z) = 0$ *eine Gleichung* n^{ten} *Grades mit reellen Koeffizienten und bedeuten* $f'(z), f''(z), \ldots, f^{(n)}(z)$ *die Abgeleiteten von* $f(z)$, *so ist die Anzahl der Wurzeln von* $f(z) = 0$, *die zwischen den zwei reellen Zahlen* A *und* B $(A < B)$ *liegen, wenn* A *und* B *keine Wurzeln von* $f(z) = 0$ *sind, gleich oder um eine gerade Zahl kleiner als die Differenz der Vorzeichenwechsel in den zwei Reihen:*

$$f(A), f'(A), \ldots, f^{(n)}(A)$$

und

$$f(B), f'(B), \ldots, f^{(n)}(B).$$

Etwaige in den zwei Reihen auftretende Nullen sind zu streichen. Jede mehrfache Wurzel, welche die Gleichung $f(z) = 0$ *im Intervall* A *bis* B *hat, ist beim Fourierschen Satz entsprechend ihrer*

Vielfachheit zu zählen. Für eine Gleichung mit lauter reellen Wurzeln gibt der Fouriersche Satz die genaue Anzahl ihrer reellen Wurzeln zwischen irgend zwei Grenzen A und B. Der Fouriersche Satz wird mit Unrecht auch nach Fouriers Zeitgenossen Budan genannt (vgl. Darboux, *Œuvres de Fourier* **2**, 310).

Satz von Laguerre (*Journ. de math.* (3) **9**, 102 (1883), (*Nouv. Ann. de math.* (2) **18**, 9 (1879), ebenda **19**, 52 (1880), *Œuvres* **1**, 6, 68 und 75, Paris 1898): *Ist A eine positive Größe und* $f(z) = 0$ *eine Gleichung mit reellen Koeffizienten, so ist die Zahl der Zeichenwechsel, welche die Reihe:*

$$\begin{aligned} f(z) &= a_0 z^n + a_1 z^{n-1} + a_2 z^{n-2} + \cdots + a_n, \\ f_1(z) &= a_0 z^{n-1} + a_1 z^{n-2} + \cdots + a_{n-1}, \\ f_2(z) &= a_0 z^{n-2} + a_1 z^{n-3} + \cdots + a_{n-2}, \\ &\vdots \\ f_n &= a_0 \end{aligned}$$

für $z = A$ *aufweist, wenigstens gleich der Zahl der Wurzeln von* $f(z) = 0$ *(mehrfache Wurzeln sind entsprechend ihrer Vielfachheit zu zählen), die größer als A sind. Übertrifft die Anzahl der Zeichenwechsel die Zahl der Wurzeln, die größer als A sind, so ist die Differenz eine gerade Zahl.*

Descartessche Zeichenregel (vgl. S. 251): *Hat man eine Gleichung* $f(z) = a_0 z^n + a_1 z^{n-1} + a_2 z^{n-2} + \cdots + a_n = 0$ *mit reellen Koeffizienten, so ist die Zahl der positiven Wurzeln von* $f(z) = 0$ *(mehrfache Wurzeln sind entsprechend ihrer Vielfachheit zu zählen) gleich oder um eine gerade Zahl kleiner als die Zahl der Vorzeichenwechsel der Gleichung. Etwa vorhandene Wurzeln* 0 *sind nicht als positive Wurzeln mitzuzählen. Unter den Zeichenwechseln bezw. Zeichenfolgen der Gleichung* $f(z) = 0$ *versteht man die Zahl der Zeichenwechsel bezw. Zeichenfolgen der Koeffizientenreihe* a_0, a_1, $a_2, \ldots, a_n$; *etwaige verschwindende Koeffizienten sind fortzulassen.*

Die Descartessche Regel ergibt sich aus dem Fourierschen Satz, indem man $A = 0$, $B = \infty$ wählt, und aus dem Laguerreschen Satz, wenn man $A = 0$ setzt. Einen Beweis der Descartesschen Regel hat Gauß, *Ges. Werke* **3**, 65, gegeben, vgl. ferner Laguerre, *Journ. de math.* (3) **9**, 99 (1883), *Œuvres* **1**, 3. Die Descartessche Regel wird irrtümlicherweise auch Harriot, *Artis analyticae praxis* (1631),

zugeschrieben. Bei Harriot findet sich die Regel überhaupt nicht (vgl. auch die Anm. in *Ostwalds Klass. d. exakt. Wiss.* Nr. 127, S. 247).

Zusätze zur Descartesschen Zeichenregel:

1. *Die Zahl der negativen Wurzeln von* $f(z) = 0$ *ist gleich oder um eine gerade Zahl kleiner als die Zahl der Vorzeichenwechsel der Gleichung* $f(-z) = 0$.

2. *Gauß' Korollar zum Descartesschen Satz* (Gauß, *Ges. Werke* **3**, 70): *Fehlen in einer Gleichung mit reellen Koeffizienten c Glieder und ist c und d die Anzahl der durch eine ungerade Zahl fehlender Glieder unterbrochenen Zeichenwechsel bezw. Zeichenfolgen in der gegebenen Gleichung, so hat die Gleichung mindestens* $e - c + d$ *imaginäre Wurzeln.*

3. Aus dem Gaußschen Korollar oder aus der am Schluß des Fourierschen Satzes gemachten Bemerkung folgt: *Eine Gleichung mit lauter reellen Wurzeln besitzt genau soviel positive Wurzeln, als* $f(z)$ *Zeichenwechsel aufweist, und soviel negative Wurzeln, wie* $f(-z)$ *Zeichenwechsel hat. Abgesehen von den letzten Gliedern können nur isolierte Glieder der Gleichung verschwinden, diese müssen stets zwischen zwei Gliedern mit verschiedenen Vorzeichen stehen.*

Eine *obere Grenze* für die Anzahl der reellen Wurzeln einer Gleichung gibt auch ein von Sylvester, *Coll. math. papers* **2**, 376, 489, 491, 493, 498, 542, 704 aufgestelltes Theorem, das eine bereits von Newton (*Arithmetica universalis* (1707), Bd. 2, Kap. 2, Maclaurin, *Algebra* (1748), Teil 2, Sect. 1, Kap. 13, Cantor, *Vorl. über Geschichte der Math.* **3**, 404 u. 561) stammende Regel als besonderen Fall enthält. Literatur: Genocchi, *Nouv. Ann. de math.* (2) **6**, 5 (1867), Petersen, *Theorie der algebraischen Gleichungen*, S. 203, Weber, *Algebra* **1**, 345, Netto, *Algebra* **1**, 225.

Die Bestimmung der Gesamtzahl der reellen Gleichungswurzeln hängt mit der für ein Intervall auf folgende Weise zusammen:

Jede Gleichung $f(z) = 0$ *mit reellen Koeffizienten hat halb so viele Wurzeln, die größer als irgendeine reelle Zahl A sind, als die Gesamtzahl reeller Wurzeln von* $f(y^2 + A) = 0$ *beträgt.* Bestimmt man also $A = 0$ entsprechend die Gesamtzahl reeller Wurzeln von $f(y^2) = 0$ nach dem Borchardtschen oder Sylvesterschen Satz, so ist die Zahl der positiven Wurzeln von $f(z) = 0$ halb so groß.

Jede Gleichung $f(z) = 0$ *mit reellen Koeffizienten hat*

zwischen den Grenzen A und B $(A < B)$ soviel Wurzeln als die Gleichung $\varphi(u) = (1+u)^n f\left(\frac{A+Bu}{1+u}\right) = 0$ *positive Wurzeln besitzt.* Die auf $\varphi(u) = 0$ angewandte Descartessche Zeichenregel gibt also analog dem Fourierschen Satz eine approximative obere Grenze für die im Intervall A bis B gelegenen Wurzeln von $f(z) = 0$ (Jacobi, *Journ. f. Math.* **13**, 348 (1835), *Ges. Werke* **3**, 279).

Die Gleichung $f(z) = 0$ hat zwischen den Grenzen A und B $(A < B)$ halb so viele Wurzeln wie $\varphi(y^2) = (1+y^2)^n f\left(\frac{A+By^2}{1+y^2}\right) = 0$ *reelle Wurzeln besitzt, und die Gesamtzahl reeller Wurzeln von $\varphi(y^2) = 0$ ist gleich der Differenz der Gesamtzahl reeller Wurzeln von $f(y^2 + A) = 0$ und der von $f(y^2 + B) = 0$.*

Über eine von Klein angegebene geometrische Vergleichung der verschiedenen Abschätzungskriterien für die Anzahl der reellen Gleichungswurzeln vgl. Weber, *Algebra* **1**, 354.

Mit Hilfe des Sturmschen Divisionsverfahrens läßt sich für eine rationale Funktion auch der von Cauchy eingeführte *Index der Funktion* bestimmen. $\varphi(t)$ sei irgendeine in dem Intervall $A \leqq t \leqq B$ (A und B reelle Zahlen) reelle und, abgesehen von ν in dem Intervall gelegenen Ausnahmestellen $t_1, t_2, \ldots, t_\nu$, eindeutige und stetige Funktion von t; die Unstetigkeitsstellen seien derartig, daß der Wert von φ mit bestimmtem Vorzeichen unendlich wird, sowohl wenn t von links her als von rechts her sich dem t_i $(i = 1, 2, \ldots, \nu)$ nähert. Für jede einzelne Unstetigkeitsstelle sei $\varepsilon(t_i) = 0, +1$ oder -1, je nachdem φ beim Überschreiten von t_i sein Vorzeichen nicht ändert oder von einem negativen zu einem positiven oder von einem positiven zu einem negativen Wert übergeht. *Unter dem Index $J_A^B(\varphi)$ der Funktion φ für das Intervall $A \leqq t \leqq B$ versteht man nach Cauchy* (*J. éc. polyt.*, Cah. **25**, 176 (1837)) *die Summe* $J_A^B(\varphi) = \varepsilon(t_1) + \varepsilon(t_2) + \cdots + \varepsilon(t_\nu)$. *Ist φ für das ganze Intervall stetig, so ist $J_A^B(\varphi) = 0$.*

Hat auch $\frac{1}{\varphi}$ ebenso wie φ für das Intervall A bis B nur einzelne im Innern des Intervalls liegende polare Unstetigkeiten, so besteht die Relation:

$$J_A^B(\varphi) + J_A^B\left(\frac{1}{\varphi}\right) = \frac{1}{2}\Big(\operatorname{sign}\varphi(B) - \operatorname{sign}\varphi(A)\Big).$$

Ist $\varphi = \frac{f_1}{f}$ *eine rationale Funktion, wobei f und f_1 teiler-*

fremde ganze Funktionen bedeuten, und ist f mindestens von dem gleichen Grade wie f_1, so findet man den Index $J_A^B\left(\frac{f_1}{f}\right)$ auf folgende Weise: Man bilde nach dem Sturmschen Divisionsverfahren die Funktionenreihe $f, f_1, f_2, \ldots, f_r$, die erhalten wird, indem man f durch f_1 dividiert und die sukzessiv auftretenden Reste immer mit entgegengesetzten Vorzeichen nimmt. $J_A^B\left(\frac{f_1}{f}\right)$ ist gleich der Anzahl der Vorzeichenwechsel, welche die Reihe

$$f(A),\ f_1(A),\ f_2(A), \ldots,\ f_r(A)$$

mehr aufweist als die Reihe

$$f(B),\ f_1(B),\ f_2(B), \ldots,\ f_r(B).$$

(Sturm in der S. 256 zitierten Arbeit Artikel 20, ferner *Journ. de math.* **1**, 305 (1836), Cauchy, a. a. O., S. 187.)

Ist f_1 von höherem Grade als f und hat man den Index $J_A^B\left(\frac{f_1}{f}\right)$ zu bestimmen, so geschieht dies mit Hilfe der oben zwischen $J_A^B(\varphi)$ und $J_A^B\left(\frac{1}{\varphi}\right)$ angegebenen Relation.

Ist f' die erste Abgeleitete von f, so ist $J_A^B\left(\frac{f'}{f}\right)$ gleich der Anzahl der zwischen A und B gelegenen reellen Wurzeln von $f(z) = 0$. (Sturmscher Satz vgl. S. 340.)

Die Anzahl der Wurzeln einer Gleichung $f(z) = 0$ mit beliebigen (auch komplexen) Koeffizienten im Innern einer einfach geschlossenen regulären Kurve C wird durch ein Cauchysches Integral bestimmt. *Ist $f(z)$ eine überall im Innern und auf dem Rande von C holomorphe und auf dem Rande auch von Null verschiedene Funktion, so ist die Anzahl ihrer innerhalb von C gelegenen Nullstellen (jede nach ihrer Vielfachheit gezählt) gleich $\frac{1}{2\pi i}\int_C \frac{f'(z)}{f(z)}\,dz$, wobei das Integral über den Rand von C in positivem Sinne zu erstrecken ist. Ist $z = x + iy$ und trennt man $f(z)$ in seinen reellen und imaginären Bestandteil, setzt also $f(x + iy) = X(x, y) + iY(x, y)$, so wird*

$$\frac{1}{2\pi i}\int \frac{f'(z)}{f(z)}\,dz = \frac{1}{2}J\left(\frac{X}{Y}\right);$$

der Index ist längs der Kurve C zu erstrecken, d. h. er wird gleich der stets geraden Zahl $p - p'$, wenn p bezw. p' die An-

zahl von Punkten auf C angibt, in denen Y verschwindet und hierbei der Quotient $\frac{X}{Y}$ *vom Negativen zum Positiven bzw. vom Positiven zum Negativen übergeht. Man kann statt X, Y auch* $-Y$, *X verwenden* (Cauchy, vgl. Sturm u. Liouville, *Journ. de math.* **1**, 278 (1836), Sturm, ebenda, S. 290). Betreffs des Cauchyschen Index vgl. ferner Hermite, *Journ. f. Math* **52**, 39 (1856), *Œuvres* **1**, 410, Hurwitz, *Math. Ann.* **46**, 273 (1895), **64**, 517 (1907).

$f(z)$ sei eine ganze rationale Funktion und C werde von λ Bogen unicursaler Kurven gebildet. Ein Bogen C_l werde durch die rationalen Funktionen $x = \varphi(t)$, $y = \psi(t)$ dargestellt, wobei die reelle Variable t das Intervall A bis B durchlaufe und auf diese Weise den Bogen C_l einmal beschreibe. Längs des Bogens C_l geht $f(z) = X(x, y) + iY(x, y)$ über in die rationale Funktion $X(\varphi, \psi) + iY(\varphi, \psi) = R(t) + iR_1(t)$; längs des Bogens C_l wird $J\left(\frac{X}{Y}\right)$ daher gleich $J_A^B\left(\frac{R(t)}{R_1(t)}\right)$ und ist als Index einer rationalen Funktion mittels des Sturmschen Divisionsverfahrens zu finden. Der Index längs der Kurve C ist die Summe der Indizes längs $C_1, C_2, \ldots, C_\lambda$. So kann man, auf rationalem Wege die Anzahl der Wurzeln innerhalb eines von Geraden und Kreisbogen begrenzten Bereiches finden. Wählt man für C einen genügend großen Kreis, dessen Zentrum der Koordinatenursprung der Gaußschen Ebene ist, so läßt sich auf die Theorie des Index ein Beweis des *Fundamentalsatzes der Algebra* gründen. Hierauf beruhen auch ihrem Gedankengang nach der erste und vierte Gaußsche Beweis des Fundamentalsatzes der Algebra (vgl. Zitat oben S. 251, siehe auch Weber, *Algebra* **1**, 333).

Die Untersuchungen über den Cauchyschen Index sind ein spezieller Fall der von Kronecker (*Monatsb. d. Berl. Akad.* (1869), (1873) u. (1878), *Ges. Werke* **1**, 175, 213, 303, **2**, 37, 113) stammenden *Charakteristikentheorie.* Sie bestimmt die mehreren Gleichungen mit reellen Koeffizienten und reellen Variablen gemeinsamen Wurzeln. Vgl. Picard, *Traité d'analyse,* 2. Aufl., Bd. **1**, 136 und **2**, 205, Weber, *Algebra* **1**, 323.

§ 16. Approximation der Wurzeln.

Jede Gleichung $b_0 + b_1 z^{\nu_1} + b_2 z^{\nu_2} + \cdots + b_k z^{\nu_k} = 0$, *bei der* $0 < \nu_1 < \nu_2 < \cdots < \nu_k$ *und* $b_0 \neq 0$, $b_1 \neq 0$ *ist, hat, wenn man ihre Wurzeln in der sog. Gaußschen Ebene der komplexer*

Zahlen darstellt, innerhalb oder am Rande eines jeden der vier Kreise mit den Radien:

$$(1)\qquad \left(\frac{\nu_2\,\nu_3\cdots\nu_k}{(\nu_2-\nu_1)\,(\nu_3-\nu_1)\cdots(\nu_k-\nu_1)}\right)^{\frac{1}{\nu_1}}\left|\frac{b_0}{b_1}\right|^{\frac{1}{\nu_1}},$$

$$(2)\qquad \left(\frac{(\nu_1+1)\,(\nu_1+2)\cdots(\nu_1+k-1)}{1\cdot 2\cdots k-1}\right)^{\frac{1}{\nu_1}}\cdot\left|\frac{b_0}{b_1}\right|^{\frac{1}{\nu_1}},$$

$$(3)\qquad k\left|\frac{b_0}{b_1}\right|^{\frac{1}{\nu_1}},$$

$$(4)\qquad kB$$

mindestens eine Wurzel. Die bei (4) *auftretende Größe B bedeutet jene der zwei positiven Zahlen* 1 *und* $\left|\frac{b_0}{b_1}\right|$, *die nicht kleiner als die andere ist* (Fejér, *Math. Ann.* **65**, 413 (1908)). Die Kreisradien hängen außer von den Anfangsgliedern b_0 und b_1 noch von der Gliederanzahl der Gleichung oder den Exponenten ab; b_0 und b_1 *allein* bestimmen keinen endlichen Kreis, der mindestens eine Gleichungswurzel enthält.

Hingegen gilt folgendes Theorem (Landau, *Sitzungsb. d. Berl. Akad.* (1904), 1118, *Vierteljahrshefte der naturf. Ges. zu Zürich* **51**, 252 (1906)): *Für jede ganze rationale (sogar ganze transzendente) Funktion* $f(z)=c_0+c_1z+c_2z^2+\cdots$, *bei der der Koeffizient* c_1 *von Null verschieden ist, existiert eine nur von den zwei Größen* c_0 *und* c_1 *abhängige positive Zahl R, so daß in einem mit dem Radius R um den Koordinatenursprung geschlagenen Kreise mindestens eine der beiden Gleichungen* $f(z)=0$ *oder* $f(z)=1$ *eine Wurzel besitzt.*

Ist $f'(z)$ *die erste Abgeleitete irgendeiner ganzen rationalen Funktion* $f(z)$, *so stellt sich in der Gaußschen Ebene der komplexen Zahlen keine der Wurzeln der Gleichung* $f'(z)=0$ *außerhalb jenes kleinsten konvexen geradlinigen Polygons dar, das sich um die Wurzeln von* $f(z)=0$ *spannen läßt* (vgl. die Literaturangaben von Fejér, *Math. Ann.* **65**, 417 (1908), ferner Cesàro, *Nouv. Ann. de math.* (3) **4**, 329 (1885), Cesàro-Kowalewski, *Elementares Lehrbuch der algebraischen Analysis* usw., S. 434).

Stellt man die Wurzeln irgendeiner Gleichung

$$f(z)=a_0z^n+a_1z^{n-1}+a_2z^{n-2}+\cdots+a_n=0$$

mit beliebigen reellen oder komplexen Koeffizienten in der sog. Gaußschen Ebene der komplexen Zahlen dar, so liegt keine von

ihnen außerhalb des mit dem Radius G um den Koordinatenursprung geschlagenen Kreises, wo G die einzige positive Wurzel der Gleichung:

$$|a_0|u^n - \{|a_1|u^{n-1} + |a_2|u^{n-2} + a_3|u^{n-3} + \ldots + |a_n|\} = 0$$

ist, d. h. G ist eine obere Grenze für den absoluten Betrag der Gleichungswurzeln (Cauchy, *Exercices de math.* (1829), *Œuvres* (2) **9**, 122). Diese Grenze ist schärfer als die von Gauß beim vierten Beweise des Fundamentalsatzes im Jahre 1849 (*Ges. Werke* **3**, 76, vgl. oben S. 251) verwendete. Aus den unten angegebenen Kriterien für die obere Grenze der positiven Wurzeln einer Gleichung mit reellen Koeffizienten folgt: Ist A die größte Zahl unter den absoluten Beträgen von $|a_1|, |a_2|, \ldots, |a_n|$, so ist die angegebene Größe G nicht größer als $1 + \frac{A}{|a_0|}$. Ist k die Anzahl der von Null verschiedenen Größen $a_1, a_2, \ldots, a_n$, so ist G nicht größer als die größte unter den Zahlen

$$\frac{k|a_1|}{|a_0|}, \quad \sqrt{\frac{k|a_2|}{|a_0|}}, \quad \sqrt[3]{\frac{k|a_3|}{|a_0|}}, \ldots, \sqrt[n]{\frac{k|a_n|}{|a_0|}}$$

Setzt man $z = \frac{1}{t}$, so erhält man für t die Gleichung $a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n = 0$; ist G' eine obere Grenze für die absoluten Beträge $|t|$ der Wurzeln der transformierten Gleichung $t^n f\left(\frac{1}{t}\right) = 0$, so ist $K = \frac{1}{G'}$ eine untere Grenze für die absoluten Beträge der Wurzeln von $f(z) = 0$, d. h. $|z| \geqq K$.

Eine Zahl G heißt eine *obere Grenze für die positiven Gleichungswurzeln, wenn alle positiven Wurzeln nicht größer als G sind;* eine solche wird durch folgende Sätze geliefert:

Ist $f(z) = a_0 z^n + a_1 z^{n-1} + a_2 z^{n-2} + \cdots + a_n = 0$ *eine Gleichung mit reellen Koeffizienten und* $a_0 > 0$, *so ist jede Zahl G, für die* $f(z)$ *und sämtliche sukzessive Abgeleitete* $f'(z), f''(z), \ldots, f^{(n-1)}(z)$ *positiv werden, eine obere Grenze für die positiven Gleichungswurzeln von* $f(z) = 0$ (Newton, *Arithmetica universalis*, Bd. 2, Kap. 4, Folge des Fourierschen Satzes auf S. 344). *Ist ebenfalls* $a_0 > 0$, *so ist auch jede Zahl G, für die alle Funktionen*

$$f_i(z) = a_0 z^{n-i} + a_1 z^{n-i-1} + \cdots a_{n-i} \quad (i = n-1, n-2, \ldots, 0)$$

positiv ausfallen, eine obere Grenze für die positiven Wurzeln (Laguerre, *Nouv. Ann. de math.* (2) **19**, 49 (1880), *Œuvres* **1**, 72, Folge des Laguerreschen Satzes auf S. 345).

Ist in der Gleichung $a_0 z^n + a_1 z^{n-1} + \cdots + a_n = 0$ *mit reellen Koeffizienten* a_j *der erste der negativen Koeffizienten, also* $a_0, a_1, \ldots, a_{j-1}$ *positiv, und ist* N *der absolute Betrag des dem absoluten Betrage nach größten negativen Koeffizienten, so ist jede der Größen*

$$1 + \sqrt[j-f]{\frac{N}{a_0 + a_1 + \cdots a_f}},$$

wobei f *jeden der Werte* $0, 1, 2, \ldots, j-1$ *annehmen kann, eine obere Grenze für die positiven Gleichungswurzeln. Im besonderen sind also*

$$1 + \sqrt[j]{\frac{N}{a_0}}, \quad 1 + \frac{N}{a_0 + a_1 + \cdots a_{j-1}}$$

(und a fortiori $1 + \frac{N}{a_0}$*, die sog. Maclaurinsche Grenze, die aber schon Rolle, Traité d'algèbre (1690), 2tes Buch, Kap. 6 hat) obere Grenzen für die positiven Gleichungswurzeln.* Vgl. auch Lagrange, *Rés. des équ. num.* (oben S. 256), *Œuvres* 8, 32, Longchamps, *Nouv. Ann. de math.* (2) **19**, 71 (1880).

Ist $a_0 > 0$ *und hat die Gleichung*

$$a_0 z^n + a_1 z^{n-1} + \cdots + a_n = 0$$

mit reellen Koeffizienten im ganzen k *negative Koeffizienten, die* a_j, a_m, a_p, a_s, ... *lauten mögen, so ist die größte der* k *Größen*

$$\sqrt[j]{k\frac{|a_j|}{a_0}}, \quad \sqrt[m]{k\frac{|a_m|}{a_0}}, \quad \sqrt[p]{k\frac{|a_p|}{a_0}}, \quad \sqrt[s]{k\frac{|a_s|}{a_0}}, \ldots$$

eine obere Grenze für die positiven Gleichungswurzeln (Cauchy, *Œuvres* (2) **9**, 152, Abel Transon, *Nouv. Ann. de math.* (2) **11**, 256 (1872)).

Ist G' eine obere Grenze für die positiven Wurzeln von $f(-z) = 0$, so ist $-G'$ eine untere Grenze für die negativen Wurzeln von $f(z) = 0$, d. h. keine der Wurzeln ist kleiner als $-G'$. Sind L und $-L'$ obere bezw. untere Grenzen für die positiven bezw. negativen Wurzeln von $z^n f\left(\frac{1}{z}\right) = 0$, so liegt keine Wurzel von $f(z) = 0$ zwischen $\frac{1}{L}$ und $-\frac{1}{L'}$.

Mittels des Sturmschen oder Fourierschen Satzes kann man zuerst die reellen Wurzeln einer Gleichung mit reellen Koeffizienten *separieren*, d. h. ein Intervall finden, in dem nur eine Wurzel der Gleichung liegt. Die zunächst folgenden Me-

thoden setzen eine auf irgendwelche Weise vorgenommene Separation der Wurzeln voraus.

Regula falsi (Regel des doppelten falschen Ansatzes, Historisches bei Cantor, Vorl. über Geschichte der Math. unter dem Stichwort *Falscher Ansatz* des Registers. vgl. auch Rudolf Wolf, *Handbuch der Astronomie* **1**, 87, Zürich 1890): *Sind A und B zwei reelle Zahlen, $f(z)$ irgendeine Funktion mit reellen Koeffizienten und sign $f(A) = -$ sign $f(B)$, so liegt*

$$B_1 = \frac{Bf(A) - Af(B)}{f(A) - f(B)} = A - \frac{(A-B)f(A)}{f(A) - f(B)} = B - \frac{(A-B)f(B)}{f(A) - f(B)}$$

stets zwischen A und B. Eine Gleichung $f(z) = 0$, die zwischen A und B eine einzige Wurzel α hat, besitzt diese daher in einem der zwei kleineren Intervalle A bis B_1 oder B_1 bis B; die Wurzel α liegt zwischen $B_1 - L$ und $B_1 + L$, wobei

$$L = \frac{|A-B|^3}{8} \cdot \frac{M}{|f(A) - f(B)|}$$

ist und M den größten Wert bedeutet, den der absolute Betrag $|f''(z)|$ der zweiten Abgeleiteten im Intervall A bis B annehmen kann (vgl. Lüroth, *Vorl. über numerisches Rechnen*, Leipzig 1900, S. 175 u. 193).

Zwischen den zwei reellen Zahlen A und B habe die Gleichung $f(z) = 0$ mit reellen Koeffizienten nur eine reelle Wurzel α, mithin ist also sign $f(A) = -$ sign $f(B)$. Das Intervall A bis B sei so klein gewählt, daß die zweite Abgeleitete $f''(z)$ in dem ganzen Intervall A bis B keine Wurzel besitzt. Dies ist möglich, nachdem $f(z)$ und $f''(z)$ vorher von etwaigen gemeinsamen Teilern befreit wurden. $f''(z)$ hat dann im ganzen Intervall A bis B dasselbe Vorzeichen. Mit A sei, unabhängig davon, ob sie die größere oder kleinere Zahl ist, diejenige der zwei Grenzen bezeichnet, für die $f(z)$ und $f''(z)$ *dasselbe Vorzeichen* haben. Auf die Grenze A kann man dann die *Newtonsche Näherungsmethode in der von Fourier verbesserten Form* anwenden (Newton (1669) vgl. oben S. 256, zuerst publiziert in Wallis' *Algebra* (1685), Lagrange, *Rés. des équ. num.*, *Œuvres* **8**, 159, Fourier, *Anal. des équ. dét.*, vgl. oben S. 256, Darboux, *Nouv. Ann. de math.* (2) **8**, 17 (1869), Fouret, ebenda (2) **9**, 567 (1890)): *Bildet man sukzessiv*

$$A_1 = A - \frac{f(A)}{f'(A)}, \quad A_2 = A_1 - \frac{f(A_1)}{f'(A_1)}, \ldots$$

$$A_{n+1} = A_n - \frac{f(A_n)}{f'(A_n)}, \ldots,$$

so wird die erste Abgeleitete $f'(z)$ im ganzen Intervall A bis α nicht Null; A_1 liegt stets zwischen A und α, A_2 zwischen A_1 und α usw., und die Größen $A, A_1, A_2, \ldots$ konvergieren nach α. Es ist α stets zwischen $A_{n+1} - L_1$ und $A_{n+1} + L_1$ gelegen; dabei ist

$$L_1 = \frac{|A_n - B|^2}{2} \frac{M_1}{|f'(A_n)|},$$

und M_1 bedeutet den größten Wert des absoluten Betrages von $f''(z)$ im Intervall A_n bis B.

Die *Kombination der Regula falsi mit der Newtonschen Näherungsmethode* (Fourier, *Anal. des équ. dét.*) ergibt unter den gleichen Bedingungen und bei Verwendung der gleichen Bezeichnungen wie beim Newtonschen Näherungsverfahren folgendes Theorem:

Bildet man sukzessiv

$$B_1 = \frac{Bf(A) - Af(B)}{f(A) - f(B)}, \quad B_{n+1} = \frac{B_n f(A_n) - A_n f(B_n)}{f(A_n) - f(B_n)} \quad (n = 1, 2, \ldots),$$

so liegt B_1 stets zwischen α und B, B_2 zwischen B_1 und α usw., und die Größen $B, B_1, B_2, \ldots$ konvergieren nach α. Die Wurzel α liegt stets zwischen A_n und B_n.

Bildet man ferner sukzessiv (Fourier a. a. O.)

$$C_1 = B - \frac{f(B)}{f'(A)}, \quad C_2 = C_1 - \frac{f(C_1)}{f'(A_1)}, \cdots C_{n+1} = C_n - \frac{f(C_n)}{f'(A_n)},$$

so liegt C_1 stets zwischen B und α, C_2 zwischen C_1 und α, C_3 zwischen C_2 und α usw., und die Zahlen $C_1, C_2, \ldots$ konvergieren nach α. Die Wurzel α liegt stets zwischen A_n und C_n.

Die Newtonsche Näherungsmethode ist nur ein Spezialfall des von Legendre-Cauchy (Cauchy, *Analyse algébrique* (1821), *Œuvres* (2) **3**, 381) stammenden Verfahrens *sukzessiver Approximation*. Aus einem geeigneten Anfangswert A ist mittels einer stetigen Funktion $\varphi(z)$ sukzessiv eine Wertreihe $A_1 = \varphi(A)$, $A_2 = \varphi(A_1), \ldots, A_{n+1} = \varphi(A_n), \ldots$ herzuleiten, so daß die Größen $A_1, A_2, \ldots$ nach einem Wert α konvergieren, der Wurzel einer vorgegebenen Gleichung $f(z) = 0$ ist. Hierzu muß $\alpha = \varphi(\alpha)$ sein. Vgl. Schroeder, *Math. Ann.* **2**, 317 (1870), Isenkrahe, ebenda **31**, 309 (1888), Netto, *Algebra* **1**, 300, Lüroth, *Vorl. über numerisches Rechnen*, S. 179, Pellet, *C. R.* **133**, 917 u. 1186 (1901), Perrin, ebenda, 1189.

Lagranges Kettenbruchmethode: Hat man eine Gleichung $f(z) = 0$ mit reellen Koeffizienten und weiß man, daß sie

λ positive Wurzeln zwischen den zwei positiven ganzen Zahlen g und $g+1$ besitzt, so setze man $z = g + \frac{1}{z_1}$. Die sich dann aus $f(z) = 0$ für z_1 ergebende Gleichung $f_1(z_1) = 0$ hat soviel positive reelle Wurzeln, die größer als 1 sind, wie $f(z) = 0$ reelle Wurzeln zwischen den Grenzen g und $g+1$ hat. Ist g_1 eine derartig gewählte ganze positive Zahl, daß $f_1(z_1) = 0$ zwischen den Grenzen g_1 und $g_1 + 1$ wenigstens eine Wurzel hat, so setze man $z_1 = g_1 + \frac{1}{z_2}$ und operiere mit der sich aus $f_1(z_1) = 0$ ergebenden Gleichung $f_2(z_2) = 0$ weiter. Auf diese Weise werden die zwischen g und $g+1$ liegenden positiven Wurzeln von $f(z) = 0$ in Kettenbrüche

$$z = g + \cfrac{1}{g_1 + \cfrac{1}{g_2 \ddots}}$$

entwickelt. Lagrange, *Rés. des équ. num.*, *Œuvres* **8**, 41, M. A. Stern, *Journ. f. Math.* **11**, 142, 277 (1834), Sturm in der S. 256 zitierten Arbeit, Art. 16.

Lagrange-Bernoullisches Verfahren: *Ist* s_ν $(\nu = 1, 2 \ldots)$ *die Summe der* ν^{ten} *Potenzen der Gleichungswurzeln, so nähert sich* $\frac{s_{\nu+1}}{s_\nu}$ *für wachsendes* ν *der ihrem absoluten Betrage nach größten Gleichungswurzel, falls eine solche existiert* (Lagrange, *Rés. des équ. num.*, *Œuvres* **8**, 170). Am einfachsten findet man die s_ν, indem man $\frac{f'(z)}{f(z)}$ nach fallenden Potenzen von z entwickelt, also $\frac{f'(z)}{f(z)} = \frac{n}{z} + \frac{s_1}{z^2} + \frac{s_2}{z^3} + \cdots$ bildet (vgl. S. 267). Dieses Verfahren ordnet sich als Spezialfall in die von Daniel Bernoulli (*Comm. Acad. Petrop.* **3**, 92 (1728)) stammende und dann von Euler (*Introductio in analysin infinitorum* (1748), Bd. 1, Kap. 17) aufgenommene Methode der Berechnung der Gleichungswurzeln mittels *rekurrenter Reihen* und läßt sich so ausgestalten (Fourier, *Anal. des équ. dét.*, *Ostwalds Klass.* Nr. 127, S. 63, Stern, *Journ. f. Math.* **11**, 293 (1834), Jacobi, ebenda **13**, 349 (1835), *Ges. Werke* **3**, 280, F. Cohn, *Math. Ann.* **44**, 473 (1894)), daß es zur Herleitung *sämtlicher* Wurzeln einer beliebigen Gleichung mit reellen oder imaginären Koeffizienten dienen kann.

Gräffesche Methode: Ist eine Gleichung $f(z) = 0$ mit den

Wurzeln $\alpha_1, \alpha_2, \ldots, \alpha_n$ gegeben und schreibt man $f(z) = g_1(z^2) + z g_2(z^2)$, faßt man also die Glieder mit geraden und ungeraden Potenzen zusammen, so hat die Gleichung

$$f_1(y) = g_1(y)^2 - y^2 g_2(y)^2 = 0$$

die Wurzeln $\alpha_1^2, \alpha_2^2, \ldots, \alpha_n^2$. Durch sukzessives Fortsetzen dieses Verfahrens findet man eine Gleichung:

$$f_\nu(u) = d_0 u^n + d_1 u^{n-1} + d_2 u^{n-2} + \cdots + d_n = 0$$

mit den Wurzeln $\alpha_1^\tau, \alpha_2^\tau, \ldots, \alpha_n^\tau$, wobei $\tau = 2^\nu$ ist. Befriedigen die absoluten Beträge der Wurzeln der vorgelegten Gleichung die Ungleichheiten $|\alpha_1| > |\alpha_2| \ldots > |\alpha_n|$, so ist für genügend große ν angenähert

$$\alpha_1^\tau = -\frac{d_1}{d_0}, \quad \alpha_2^\tau = -\frac{d_2}{d_1}, \quad \alpha_3^\tau = -\frac{d_3}{d_2}, \cdots, \alpha_n^\tau = -\frac{d_n}{d_{n-1}},$$

und man findet $\alpha_1, \alpha_2, \ldots, \alpha_n$ durch Ziehen τ^{ter} Wurzeln.

Ist

$$|\alpha_1| = |\alpha_2| \cdots = |\alpha_j|$$

aber größer als

$$|\alpha_{j+1}|, |\alpha_{j+2}|, \ldots, |\alpha_n|,$$

so befriedigen die Größen $\alpha_1^\tau, \alpha_2^\tau, \ldots, \alpha_j^\tau$ für genügend große τ angenähert die Gleichung $d_0 u^j + d_1 u^{j-1} \cdots + d_j = 0$ und die Größen $\alpha_{j+1}^\tau, \alpha_{j+2}^\tau, \ldots, \alpha_n^\tau$ die Gleichung

$$d_j u^{n-j} + d_{j+1} u^{n-j-1} + d_{j+2} u^{n-j-2} + \cdots + d_n = 0.$$

Man kann die Methode so ausgestalten, daß sie *alle* Gleichungswurzeln liefert. Gräffe, *Die Auflösung der höheren numerischen Gleichungen*, Zürich 1837. Encke, *Journ. f. Math.* **22**, 193 (1841), Carvallo, *Ann. de la faculté de Toulouse* **3** (1889) (*Thèse de Paris*).

Wir schließen mit der Aufsuchung der *rationalen Wurzeln einer Gleichung mit rationalen Koeffizienten:*

Die rationalen Wurzeln einer Gleichung mit ganzzahligen Koeffizienten, bei der der Koeffizient der höchsten Potenz gleich 1 *ist, müssen ganzzahlig sein.*

Die rationalen Wurzeln einer Gleichung

$$a_0 z^n + a_1 z^{n-1} + a_2 z^{n-2} + \cdots + a_n = 0$$

mit ganzzahligen Koeffizienten findet man, indem man $a_0 z = y$ *setzt und die ganzzahligen Wurzeln von*

$$y^n + a_1 y^{n-1} + a_2 a_0 y^{n-2} + a_3 a_0^2 y^{n-3} + \cdots + a_n a_0^{n-1} = 0$$

sucht (Rolle, *Traité d'algèbre* (1690), 2tes Buch, Kap. 5).

Man findet die etwa vorhandenen ganzzahligen Wurzeln der Gleichung

$$z^n + a_1 z^{n-1} + \cdots + a_n = 0$$

mit ganzzahligen Koeffizienten, indem man alle ganzzahligen positiven und negativen Teiler von a_n *sucht, die innerhalb der Wurzelgrenzen liegen. Ist* α *ein solcher Wert, so ist* α *dann und nur dann Gleichungswurzel, wenn sämtliche Größen*

$$\frac{a_n}{\alpha} = b_1, \quad \frac{b_1 + a_{n-1}}{\alpha} = b_2, \quad \frac{b_2 + a_{n-2}}{\alpha} = b_3, \ldots,$$

$$\frac{b_{n-2} + a_2}{\alpha} = b_{n-1}$$

ganzzahlig ausfallen und schließlich $\frac{b_{n-1} + a_1}{\alpha} = -1$ *wird.*

Kapitel V.

Invariantentheorie.

Von *H. E. Timerding* in Braunschweig.

§ 1. Binärformen zweiter bis vierter Ordnung.

Die Invariantentheorie beschäftigt sich mit den ganzen rationalen Funktionen, insbesondere den homogenen Funktionen oder Formen, bei denen alle Glieder bezüglich der veränderlichen Größen von derselben Dimension sind. Doch sind auf die letzteren sofort die nicht homogenen Funktionen zurückzuführen, indem man eine der Variablen $= 1$ annimmt. Die Invariantentheorie behandelt diese Formen aber nicht allgemein, sondern nur hinsichtlich ihres Verhaltens bei homogenen linearen Transformationen der Veränderlichen. *Invariante* heißt nach Sylvester (*Cambr. Dublin Math. Journ.* **6** (1851) 290) eine solche Funktion der Koeffizienten einer Form, die für die transformierte Form gebildet bis auf einen konstanten (d. h. von den Koeffizienten der Form unabhängigen) Faktor mit dem analogen, für die ursprüngliche Form gebildeten Ausdrucke übereinstimmt. Der konstante Faktor ist hierbei eine Potenz der Determinante oder des *„Moduls"* jener linearen Substitution, durch welche die Transformation der Variablen gegeben wird. Der Exponent dieser Potenz heißt das *Gewicht* der Invariante, die Dimension, zu welcher in *jedem* ihrer Glieder die Koeffizienten der Form vorkommen, der *Grad* der Invariante. Zur Unterscheidung wird die Dimension, zu welcher in einer Form die Veränderlichen vorkommen, immer als *Ordnung* bezeichnet.

Das einfachste Beispiel wird durch eine quadratische Form zweier Veränderlichen oder *Binärform zweiter Ordnung:*

$$f = a_0 x_1^2 + 2 a_1 x_1 x_2 + a_2 x_2^2$$

gegeben. Geht diese durch die lineare Transformation:

$$x_1 = \alpha x_1' + \beta x_2', \; x_2 = \gamma x_1' + \delta x_2'$$

in die Form:

$$f = a_0' x_1'^2 + 2 a_1' x_1' x_2' + a_2' x_2'^2$$

über, so wird, wenn

$$\Delta = \alpha\delta - \beta\gamma$$

den *Modul* der Substitution bezeichnet,

$$a_0' a_2' - a_1'^2 = \Delta^2 (a_0 a_2 - a_1^2).$$

Die *„Diskriminante"*

$$D = 2(a_0 a_2 - a_1^2)$$

hat also die Invarianteneigenschaft, und die letztere ist an diesem Beispiel zuerst durch Gauß (*Disquisitiones arithmeticae* 1801, *Werke* **1**, Art. 157) konstatiert worden.

Der Diskriminante D einer quadratischen Form steht die *Resultante R zweier linearen Formen*

$$u = a_1 x_1 + a_2 x_2, \; v = b_1 x_1 + b_2 x_2$$

zur Seite. Es ist

$$R = a_1 b_2 - a_2 b_1.$$

Auch diese Diskriminante scheidet nur einen Faktor, nämlich das Quadrat des Substitutionsmoduls aus, wenn u und v durch dieselbe lineare Substitution transformiert werden, und heißt deshalb eine *simultane Invariante* von u und v.

Das System *zweier quadratischer Binärformen* ist zuerst von Boole (*Cambr. Math. Journ.* **3** (1843) 11 ff.) behandelt worden. Sind

$$f = a_0 x_1^2 + 2 a_1 x_1 x_2 + a_2 x_2^2, \; g = b_0 x_1^2 + 2 b_1 x_1 x_2 + b_2 x_2^2$$

die beiden Formen, so kann man aus ihnen eine Form

$$f_\lambda = f + \lambda g = (a_0 + \lambda b_0) x_1^2 + 2(a_1 + \lambda b_1) x_1 x_2 + (a_2 + \lambda b_2) x_2^2$$

zusammensetzen. Die Diskriminante:

$$D_\lambda = 2\{(a_0 + \lambda b_0)(a_2 + \lambda b_2) - (a_1 + \lambda b_1)^2\}$$

dieser Form läßt sich in folgender Gestalt schreiben:

$$D_\lambda = D_{11} + 2 D_{12} \lambda + D_{22} \lambda^2.$$

Hierbei sind D_{11} und D_{22} die Diskriminanten von f und g, und es wird:

$$D_{12} = a_0 b_2 - 2 a_1 b_1 + a_2 b_0.$$

Da D_1, D_{11}, D_{22} die Invarianteneigenschaft haben, muß sie auch D_{12} haben. D_{12} ist eine *simultane Invariante* der Formen f, g. Eine andere solche simultane Invariante ist die folgende:

$$R = D_{11} D_{22} - D_{12}^2$$

oder

$$R = 4(a_0 b_1 - a_1 b_0)(a_1 b_2 - a_2 b_1) - (a_0 b_2 - a_2 b_0)^2.$$

Diese ist die doppelte Diskriminante der aus f, g ableitbaren quadratischen Form:

$$H = (a_0 b_1 - a_1 b_0) x_1^2 + (a_0 b_2 - a_2 b_0) x_1 x_2 + (a_1 b_2 - a_2 b_1) x_2^2.$$

Diese Form läßt sich auch schreiben:

$$H = (a_0 x_1 + a_1 x_2)(b_1 x_1 + b_2 x_2) - (a_1 x_1 + a_2 x_2)(b_0 x_1 + b_1 x_2),$$

oder mit Benutzung von Differentiationssymbolen:

$$H = \frac{1}{4} \begin{vmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} \\ \frac{\partial g}{\partial x_1} & \frac{\partial g}{\partial x_2} \end{vmatrix}.$$

Eine solche aus den partiellen Derivierten mehrerer Formen gebildete Determinante heißt *Funktional-* oder *Jacobische Determinante* (vgl. S. 156). Transformiert man f, g durch homogene lineare Substitutionen und bildet für die transformierten Formen die Form H, so unterscheidet sich der gefundene Wert von dem obenstehenden nur um eine Potenz (nämlich den Kubus) der Substitutionsdeterminante Δ. Eine Form von dieser Eigenschaft heißt, ebenfalls nach Sylvester, eine (simultane) *Kovariante* der Formen f, g. $R = 0$ ist die Bedingung dafür, daß H das Quadrat einer linearen Form wird, und gleichzeitig dafür, daß die Gleichungen $f = 0$, $g = 0$ eine gemeinsame Wurzel $x_1 : x_2$ besitzen. R heißt infolge dieser letzten Eigenschaft die Resultante von f und g. Die drei Formen f, g, H genügen der identischen Beziehung:

$$H^2 = -\tfrac{1}{2}(D_{11} g^2 - 2 D_{12} f g + D_{22} f^2).$$

Nimmt man zu den beiden Formen f, g eine *dritte quadratische Form:*

$$h = c_0 x_1^2 + 2 c_1 x_1 x_2 + c_2 x_2^2$$

hinzu, so wird die Determinante

$$D_{123} = \begin{vmatrix} a_0 & a_1 & a_2 \\ b_0 & b_1 & b_2 \\ c_0 & c_1 & c_2 \end{vmatrix}$$

eine simultane Invariante (vom Gewicht 3) der drei Formen f, g, h. Nennt man D_{33} die Diskriminante von h und $D_{13} = D_{31}$, $D_{23} = D_{32}$ die zu $D_{12} = D_{21}$ analogen, für f, g zusammen mit h gebildeten Invarianten, so wird:

$$2 D_{123}^2 = \begin{vmatrix} D_{11} & D_{12} & D_{13} \\ D_{21} & D_{22} & D_{23} \\ D_{31} & D_{32} & D_{33} \end{vmatrix}.$$

Ferner gilt die identische Gleichung:

$$\begin{vmatrix} D_{11} & D_{12} & D_{13} & f \\ D_{21} & D_{22} & D_{23} & g \\ D_{31} & D_{32} & D_{33} & h \\ f & g & h & 0 \end{vmatrix} = 0.$$

Bildet man auch die zu H analogen Formen G, F für f, g zusammen mit h, so ergeben sich die weiteren Identitäten:

$$2 D_{123} \cdot f = D_{11} F + D_{12} G + D_{13} H,$$
$$2 D_{123} \cdot g = D_{21} F + D_{22} G + D_{23} H,$$
$$2 D_{123} \cdot h = D_{31} F + D_{32} G + D_{33} H,$$
$$F f + G g + H h = 0.$$

Die Theorie der *Binärform dritter Ordnung* und die daran sich knüpfende Auflösung der Gleichungen dritten Grades ist von Cayley (*A fifth memoir upon Quantics*, *Philos. Trans.* **148** (1858) 429, *Papers* **2**, 527, Art. 115—127) gegeben worden. Aus der Form:

$$f = a_0 x_1^3 + 3 a_1 x_1^2 x_2 + 3 a_2 x_1 x_2^2 + a_3 x_2^2$$

lassen sich durch Differentiation zwei quadratische Formen

$$f_1 = \frac{1}{3}\frac{\partial f}{\partial x_1}, \quad f_2 = \frac{1}{3}\frac{\partial f}{\partial x_2}$$

herleiten, die ausgeschrieben lauten:

$$f_1 = a_0 x_1^2 + 2a_1 x_1 x_2 + a_2 x_2^2,$$
$$f_2 = a_1 x_1^2 + 2a_2 x_1 x_2 + a_3 x_2^2.$$

Bildet man deren simultane Kovariante:

$$H = (a_0 a_2 - a_1^2) x_1^2 + (a_0 a_3 - a_1 a_2) x_1 x_2 + (a_1 a_3 - a_2^2) x_2^2,$$

so ist dies eine Kovariante, die sogenannte *Hessesche Determinante* der Grundform f (vgl. S. 160). Man kann sie schreiben:

$$H = \frac{1}{36}\left\{\frac{\partial^2 f}{\partial x_1^2}\cdot\frac{\partial^2 f}{\partial x_2^2} - \left(\frac{\partial^2 f}{\partial x_1 \partial x_2}\right)^2\right\}.$$

Die Diskriminante von H:

$$R = 4(a_0 a_2 - a_1^2)(a_1 a_3 - a_2^2) - (a_0 a_3 - a_1 a_2)^2$$

ist die einzig existierende Invariante von f. Setzen wir noch:

$$H_1 = \frac{1}{2}\frac{\partial H}{\partial x_1}, \quad H_2 = \frac{1}{2}\frac{\partial H}{\partial x_2},$$

so wird

$$Q = f_1 H_2 - f_2 H_1$$

eine neue Kovariante von f, die ausgerechnet lautet:

$$\begin{aligned} Q = {} & (a_0^2 a_3 - 3a_0 a_1 a_2 + 2a_1^3) x_1^3 \\ & + 3(a_0 a_1 a_3 - 2a_0 a_2^2 + a_1^2 a_2) x_1^2 x_2 \\ & - 3(a_0 a_2 a_3 - 2a_1^2 a_3 + a_1 a_2^2) x_1 x_2^2 \\ & - (a_0 a_3^2 - 3a_1 a_2 a_3 + 2a_2^3) x_2^3. \end{aligned}$$

Sie ist also von der dritten Ordnung. Die beiden Kovarianten H, Q und die Invariante R sind mit der Grundform f durch die Identität verknüpft:

$$Q^2 + 4H^3 + Rf^2 = 0,$$

die als die *Cayleysche Gleichung* bezeichnet wird.

Bildet man die Form:

$$= Q + \lambda f,$$

so werden deren Kovarianten, wenn man $R + \lambda^2 = \Theta$ setzt,

$$H_\lambda = \Theta \cdot H, \qquad Q_\lambda = \Theta \cdot (\lambda Q - Rf),$$

und die Invariante wird

$$R_\lambda = \Theta^2 R.$$

Unter den Formen f_λ sind zwei enthalten, welche dritte Potenzen linearer Formen u, v sind. Diese findet man für $\lambda = \pm\sqrt{-R}$, wie man sofort sieht, wenn man die Cayleysche Identität in der Form

$$(Q + \sqrt{-R} \cdot f)(Q - \sqrt{-R} \cdot f) = -4H^3$$

schreibt. Aus $Q + \sqrt{-R} \cdot f = u^3$, $Q - \sqrt{-R} \cdot f = v^3$ folgt aber

$$f = \frac{1}{2\sqrt{-R}}(u^3 - v^3) = \frac{1}{2\sqrt{-R}}(u - v)(u - \varepsilon v)(u - \varepsilon^2 v),$$

wenn ε eine primitive dritte Einheitswurzel ist. Auf diese Weise ergibt sich die Zerlegung der Binärform dritter Ordnung in Linearfaktoren und damit die Auflösung der Gleichungen dritten Grades. —

Eine *Binärform vierter Ordnung:*

$$f = a_0 x_1^4 + 4a_1 x_1^3 x_2 + 6a_2 x_1^2 x_2^2 + 4a_3 x_1 x_2^3 + a_4 x_2^4$$

hat zwei Invarianten

$$i = a_0 a_4 - 4a_1 a_3 + 3a_2^2,$$

$$j = \begin{vmatrix} a_0 & a_1 & a_2 \\ a_1 & a_2 & a_3 \\ a_2 & a_3 & a_4 \end{vmatrix}.$$

Eine Kovariante, die wieder als die Hessesche Determinante zu bezeichnen ist, finden wir in dem Ausdrucke:

$$H = \left(\frac{1}{4 \cdot 3}\right)^2 \left\{ \frac{\partial^2 f}{\partial x_1^2} \cdot \frac{\partial^2 f}{\partial x_2^2} - \left(\frac{\partial^2 f}{\partial x_1 \partial x_2}\right)^2 \right\}.$$

Dieselbe ist von der vierten Ordnung. Eine zweite Kovariante wird durch die Funktionaldeterminante von f und H

$$T = \frac{1}{8}\left(\frac{\partial f}{\partial x_1} \cdot \frac{\partial H}{\partial x_2} - \frac{\partial f}{\partial x_2} \cdot \frac{\partial H}{\partial x_1}\right)$$

gegeben. Es gilt dann die Identität:

$$T^2 = -\{4H^3 - iHf^2 + jf^3\}.$$

Diese geht, wenn man die Wurzeln der Gleichung:

$$4z^3 - iz - j = 0,$$

welche die Resolvente der Gleichung $f = 0$ heißt, mit m_1, m_2, m_3 bezeichnet, in die Form über:

$$T^2 = -4(H + m_1 f)(H + m_2 f)(H + m_3 f).$$

Auf der rechten Seite muß jeder Faktor ein vollständiges Quadrat sein, weil, besondere Ausnahmefälle abgerechnet, jeder Faktor verschieden und die linke Seite ein Quadrat ist, also läßt sich setzen:

$$H + m_1 f = -\varphi^2,\ H + m_2 f = -\psi^2,\ H + m_3 f = -\chi^2,$$

so daß

$$(m_2 - m_3)\varphi^2 + (m_3 - m_1)\psi^2 + (m_1 - m_2)\chi^2 \equiv 0$$

und

$$T = 2\varphi\psi\chi$$

wird. Aus diesen Formeln folgt die Zerlegung der Form f in Linearfaktoren. Es wird nämlich:

$$f = A u_1 u_2 u_3 u_4,$$

wobei A eine Konstante bedeutet und

$$u_1 = (m_2 - m_3)\frac{\partial\varphi}{\partial x_2} + (m_3 - m_1)\frac{\partial\psi}{\partial x_2} + (m_1 - m_2)\frac{\partial\chi}{\partial x_2},$$

$$u_2 = (m_2 - m_3)\frac{\partial\varphi}{\partial x_2} - (m_3 - m_1)\frac{\partial\psi}{\partial x_2} - (m_1 - m_2)\frac{\partial\chi}{\partial x_2},$$

$$u_3 = -(m_2 - m_3)\frac{\partial\varphi}{\partial x_2} + (m_3 - m_1)\frac{\partial\psi}{\partial x_2} - (m_1 - m_2)\frac{\partial\chi}{\partial x_2},$$

$$u_4 = -(m_2 - m_3)\frac{\partial\varphi}{\partial x_2} - (m_3 - m_1)\frac{\partial\psi}{\partial x_2} + (m_1 - m_2)\frac{\partial\chi}{\partial x_2}$$

setzt werden kann. Diese Methode der Zerfällung einer Funkon vierter Ordnung in Linearfaktoren sowie die ganze heorie der Formen vierter Ordnung geht auf Cayley zurück (a. a. O. Art. 128 ff.). Für die letzten Formeln vgl. man Gordans *Vorlesungen über Invariantentheorie* 2 (1887), 195.

Setzt man:

$$\varphi = (r_1 x_1 + r_2 x_2) \cdot (s_1 x_1 + s_2 x_2) = \xi_1 \cdot \xi_2$$

so wird:

$$\psi = \alpha(\xi_1^2 - \xi_2^2), \quad \chi = \beta(\xi_1^2 + \xi_2^2).$$

Durch die lineare Substitution:

$$\xi_1 = r_1 x_1 + r_2 x_2, \quad \xi_2 = s_1 x_1 + s_2 x_2$$

wird somit die Form f, da $(m_2 - m_3) f = \psi^2 - \chi^2$ ist, auf die *„kanonische"* Gestalt:

$$f = \gamma(\xi_1^4 + 6 m \xi_1^2 \xi_2^2 + \xi_2^4)$$

gebracht. Hierbei ist m durch die Gleichung bestimmt:

$$\frac{(1 + 3m^2)^3}{m^2(1 - m^2)^2} = \frac{i^3}{j^2}$$

Bildet man die zusammengesetzte Form

$$f_{(\lambda)} = \lambda_1 f + \lambda_2 H,$$

setzt:

$$4\Omega = 4\lambda_1^3 - i\lambda_1\lambda_2^2 - j\lambda_2^3$$

und bezeichnet die Kovarianten dieser Form dritter Ordnung mit H_Ω, Q_Ω, so werden die Invarianten von $f_{(\lambda)}$:

$$i_\lambda = -12 H_\Omega, \quad j_\lambda = -4 Q_\Omega,$$

und die Kovarianten:

$$H_\lambda = \frac{1}{3}\left(\frac{\partial \Omega}{\partial \lambda_1} H - \frac{\partial \Omega}{\partial \lambda_2} f\right),$$

$$T_\lambda = \Omega \cdot T.$$

Die Invariante von Ω ist bis auf einen Zahlfaktor:

$$R = i^3 - j^2.$$

Dies ist die Diskriminante der Form vierter Ordnung f.

Ist $R = 0$, so hat f einen doppelten Linearfaktor, ebenso H, T aber einen fünffachen Faktor.

Wenn die drei Wurzeln m_1, m_2, m_3 der kubischen Resolvente zusammenfallen, d. h. $i = 0$ und $j = 0$ wird, hat f einen dreifachen Linearfaktor, H wird, von einem konstanten

Faktor abgesehen, die vierte Potenz und T die sechste Potenz des dreifachen Linearfaktors von f.

Wenn H von f nur um einen konstanten Faktor verschieden ist, so wird f und damit H das Quadrat einer quadratischen Form. Dann wird T der dritten Potenz dieser quadratischen Form proportional.

Wenn H identisch verschwindet, ist f die vierte Potenz einer linearen Form. Dann verschwindet auch T identisch.

Die Binärformen vierter Ordnung findet man sehr vollständig behandelt bei Clebsch, *Theorie der binären algebraischen Formen* (1872), S. 134—178.

§ 2. Binärformen beliebig hoher Ordnung.

Die Untersuchung der invarianten Bildungen einer allgemeinen Binärform n^{ter} Ordnung hat sich in drei Etappen vollzogen. Die erste ist durch die Namen Sylvester und Cayley gekennzeichnet und vor allem in Sylvesters Abhandlungen *On the Principles of the Calculus of Forms, Cambr. Dubl. Math. J.* **6** (1851), 186, 289, **7** (1852), 52. 179, **8** (1853), 62, 256, **9** (1854), 85 und in Cayleys *Memoirs upon Quantics* entwickelt, die in den *Philos. Transactions* von 1854 ab erschienen und mit Zusätzen des Verfassers abgedruckt sind in den *Collected papers* I 221, II 250, III 310, IV 513, V 527, VI 561, VII 325, VIII 147, IX 334, X 339. Den Arbeiten Sylvesters und Cayleys voraufgegangen sind die bedeutungsvollen, aber wenig bekannten Aufsätze von Boole, *Cambr. Math. J.* **3** (1843) 6, 106. Eine zusammenfassende Darstellung der ganzen Theorie findet man in G. Salmons *Lessons introductory to the modern higher algebra*, 1. Aufl. 1859. (Deutsch von W. Fiedler als *Algebra der linearen Transformationen.*) Selbständige Bedeutung kommt Brioschis *Teoria dei Covarianti* (Rom 1861) zu.

Die zweite Phase der Entwicklung bildet die Theorie von Clebsch und Gordan, die man in folgenden Werken findet: A. Clebsch, *Theorie der binären algebraischen Formen,* Leipzig 1872, Faà di Bruno, *Einleitung in die Theorie der binären Formen,* Leipzig 1891, P. Gordans *Vorlesungen über Invariantentheorie,* herausgegeben von G. Kerschensteiner, Leipzig 1887.

Die dritte Stufe ist bezeichnet durch die Arbeiten von D. Hilbert, *Math. Annalen* **36** (1890), 473 und **42** (1893), 313. (Vgl. das kurze Referat von H. S. White, *Bull. of the Americ. Math. Soc.* (2) **5** (1899), 161 und über die allgemeine

Entwicklung der Invariantentheorie den Bericht von Fr. Meyer, *Math. Ver.* **1** (1892) und dessen Referat *Enzykl.* I[1] p. 320.) Die Cayley-Sylvestersche Theorie hat sich indes durch die letzten Arbeiten Sylvesters, die neue Bahnen und Ziele eröffneten, bis in die Gegenwart weiter entwickelt. Die neuere englische Invariantentheorie kann man aus dem Buche von J. H. Grace und A. Young, *The Algebra of Invariants*, Cambridge 1902, kennen lernen.

Die Binärform n^{ter} Ordnung schreibt man gewöhnlich in der Gestalt:

$$f = a_0 x_1^n + \binom{n}{1} a_1 x_1^{n-1} x_2 + \binom{n}{2} a_2 x_1^{n-2} x_2^2 + \cdots + a_n x_2^n,$$

wobei $\binom{n}{1}, \binom{n}{2} \cdots$ die Binomialkoeffizienten bezeichnen. Es sind dann die Bildungsgesetze zu suchen, welche die Invarianten und Kovarianten dieser Form f befolgen. Für die Invarianten und Kovarianten hat man die gemeinsame Bezeichnung *Komitanten*, die K. Reuschle (*Verh. d. Züricher Kongresses* 1898, S. 123) an Stelle des von Sylvester (*Cambr. Dubl. Math. J.* **6** (1851), 290) gebrauchten Ausdruckes Konkomitanten eingeführt hat.

Jede Kovariante ist eine homogene Funktion nicht bloß von den Veränderlichen x_1, x_2, sondern auch von den Koeffizienten $a_0, a_1 \ldots$ Demgemäß unterscheiden wir an ihr, wie bereits angegeben, ihre *Ordnung* h als die Dimension, in der die ersteren vorkommen, von dem *Grad* k als die Dimension, in der sie die letzteren enthält. Invarianten lassen sich als Kovarianten von der Ordnung 0 auffassen. Der Koeffizient eines Gliedes der Kovariante besteht aus Ausdrücken von folgender Form:

$$c \cdot a_0^{\lambda_0} a_1^{\lambda_1} a_2^{\lambda_2} \cdots a_n^{\lambda_n},$$

in denen c einen ganzzahligen Faktor bezeichnet und immer

$$\lambda_0 + \lambda_1 + \lambda_2 + \cdots + \lambda_n = k$$

ist. Innerhalb eines Koeffizienten hat auch für alle Glieder des ihn darstellenden Aggregates

$$\lambda_1 + 2\lambda_2 + \cdots + n\lambda_n = w$$

denselben Wert, der das *Gewicht* des betreffenden Koeffizienten heißt. Man drückt dies so aus, daß man sagt, die Koeffizienten der Kovarianten und insbesondere die Invarianten seien *homogene* und *isobare* Funktionen der Koeffizienten a_i. Bei dem

ersten Koeffizienten, nämlich dem von x_1^h ist das Gewicht $w = \frac{1}{2}(nk - h)$ und steigt bei jedem folgenden Koeffizienten um 1, so daß es bei dem letzten, nämlich dem von x_2^n, $\frac{1}{2}(nk + h)$ beträgt. Die erste Zahl w heißt das *Gewicht der Kovariante*. Es ist der Exponent der Potenz des Substitutionsmoduls, mit der sich bei einer linearen Transformation die Kovariante multipliziert, wie man sofort sieht, wenn man die besondere Substitution $x_1 = x_1'$, $x_2 = \Delta x_2'$ wählt. Für eine Invariante wird (da $h = 0$) das Gewicht $w = \frac{1}{2} ng$. Das Gewicht ist eingeführt worden von Cayley (*Philos. Trans.* **146** (1856) 101, *Papers* **2**, 250).

Die Komitanten einer Kovariante sind wieder Komitanten der Urform, und eine ganze rationale Funktion von Komitanten ist wieder eine Komitante, wenn alle ihre Glieder von derselben Ordnung, demselben Grade und demselben Gewichte sind.

Jede Kovariante Q genügt bestimmten *Differentialgleichungen*, die von Cayley (*Journ. f. Math.* **47** (1854), 109, *Papers* **2**, 164) angegeben sind, die aber Aronhold schon 1850 besaß, wie aus einem von Lampe (*Arch. d. Math.* (3) **1** (1901), 38) veröffentlichten Briefe hervorgeht. Diese Differentialgleichungen lauten:

$$\sum_{\varrho} (n - \varrho) \cdot a_\varrho \cdot \frac{\partial Q}{\partial a_\varrho} - x_1 \frac{\partial Q}{\partial x_1} = wQ,$$

$$\sum_{\varrho} \varrho \cdot a_\varrho \cdot \frac{\partial Q}{\partial a_\varrho} - x_2 \frac{\partial Q}{\partial x_2} = wQ,$$

$$\sum_{\varrho} (n - \varrho) a_{\varrho+1} \frac{\partial Q}{\partial a_\varrho} - x_1 \frac{\partial Q}{\partial x_2} = 0,$$

$$\sum_{\varrho} \varrho \cdot a_{\varrho-1} \frac{\partial Q}{\partial a_\varrho} - x_2 \frac{\partial Q}{\partial x_1} = 0.$$

Aus diesen Differentialgleichungen ist sofort zu erschließen, daß in jeder Komitante *alle* Koeffizienten der Urform vorkommen müssen. Für Invarianten ergeben sich leicht ersichtliche vereinfachte Gleichungen, indem die Derivierten nach den x verschwinden.

Die vorletzte Gleichung gestattet, aus dem ersten Gliede der Kovariante, nämlich dem mit x_1^h, sukzessive die Glieder mit $x_1^{h-1}x_2$, $x_1^{h-2}x_2^2 \cdots$, also die ganze Kovariante abzuleiten. Setzt man nämlich

$$Q = \sum_i \binom{h}{i} A_i x_1^{h-i} x_2^i,$$

so ergibt sich:

$$A_{i+1} = \frac{1}{h-i} \sum_{\varrho} (n-\varrho) a_{\varrho+1} \frac{\partial A_i}{\partial a_\varrho}.$$

Die Kovariante ist also völlig bestimmt durch ihr erstes Glied, welches das *Leitglied* heißt. Dies Leitglied A_0 genügt selbst der folgenden Differentialgleichung:

$$\sum_{\varrho} \varrho a_{\varrho-1} \frac{\partial A_0}{\partial a_\varrho} = 0,$$

und jede homogene, isobare Funktion der a_i, die dieser Differentialgleichung genügt, läßt sich als das Leitglied einer Kovariante deuten.

Diese Leitglieder werden nach Sylvester als *Seminvarianten* bezeichnet. Insbesondere heißen Seminvarianten, die nicht weiter auf andere Seminvarianten reduziert werden können, *Perpetuanten*. Z. B. erhält man bei einer Form vierter Ordnung

$$f = a_0 x_1^4 + 4 a_1 x_1^3 x_2 + 6 a_2 x_1^2 x_2^2 + 4 a_3 x_1 x_2^3 + a_4 x_2^4$$

die Seminvarianten:

$$S_0 = a_0, \qquad S_1 = \begin{vmatrix} a_0 & a_1 \\ a_1 & a_2 \end{vmatrix} \qquad S_4 = \begin{vmatrix} a_0 & a_1 & a_2 \\ a_1 & a_2 & a_3 \\ a_2 & a_3 & a_4 \end{vmatrix},$$

$$S_2 = a_0^2 a_3 - 3 a_0 a_1 a_2 + 2 a_1^3, \quad S_3 = a_0 a_4 - 4 a_1 a_3 + 3 a_2^2$$

Zwischen diesen 5 Semivarianten besteht die Beziehung:

$$4 S_1^3 + S_2^2 - S_0^2 S_2 S_3 + S_0^3 S_4 = 0.$$

Alle *Syzygien* oder identischen Relationen zwischen den Komitanten lassen sich auf solche Gleichungen zwischen den Seminvarianten reduzieren. S_0, S_1, S_2, S_3 heißen nach Sylvester die „*Grundformen*".

Führt man die Seminvarianten an Stelle der Koeffizienten a_0, a_1 ... in die obigen Differentialgleichungen ein, so reduzieren sich diese auf *drei* Gleichungen, indem an Stelle der beiden letzten eine einzige Differentialgleichung tritt, deren Integrale sämtliche *Kovarianten* (bzw. *Invarianten*) der Urform sind. Vgl. Junker, *Math. Ann.* **64** (1907), 328.

Die Theorie der Seminvarianten ist entwickelt worden von Sylvester, *Am. Journ. of Math.* **5** (1882), 79, **6** (1883), 97, Cayley, ebenda **7** (1884), 1, 59, **15** (1893), 1 (*Papers* **12**,

239, 275, **13**, 264), Hammond, ebenda **8** (1886), 104, Perrin, *Bull. Soc. math. de France* **11** (1883), 88, Mac Mahon, *Am. Journ. of Math.* **7** (1884), 26, 259, **8** (1885), 1, d'Ocagne, *Comptes Rendus* **102** (1886), 916, **104** (1887), 961, 1364, *Bruxelles Ann. Soc. scient.* $\mathbf{10^B}$ (1886), 75, $\mathbf{11^B}$ (1887), 314, $\mathbf{12^B}$ (1888), 185, Roberts, *Proc. Lond. Math. Soc.* **21** (1889), 113. Eine zusammenhängende Darstellung gab Deruyts, *Théorie générale des formes algébriques* (Brüssel 1891). Ergänzungen dazu gab der Verfasser in den Schriften der Brüsseler Akademie: *Bull.* (3) **26** (1893), 258, **28** (1894), 157; *Mém.* **51** (1893), **52** (1894).

Man kann aus dem Leitglied A_0 die Kovariante Q auch auf folgende Art herleiten, die von Bruno (*Journ. f. Math.* **90** (1880), 186) herrührt. Man schreibe die Form f inhomogen, indem man $x_1 = x$, $x_2 = 1$ setzt und bilde sukzessive die Funktionen

$$f_0 = f, \quad f_1 = \frac{1}{n}\frac{df_0}{dx}, \quad f_2 = \frac{1}{n-1}\frac{df_1}{dx} \cdots$$

Wenn man dann in A_0

$$f_0, f_1, f_2 \cdots \text{ für } a_0, a_1, a_2 \cdots$$

schreibt, so geht A_0 in Q über. Daraus folgt mit Rücksicht auf die Differentialgleichung, der A_0 genügt, daß jede isobare, homogene Funktion Q von $f_0, f_1, f_2 \cdots$, die der Bedingung genügt:

$$\sum_{\varrho} \varrho f_{\varrho-1} \frac{\partial Q}{\partial f_\varrho} = 0,$$

als Funktion von x betrachtet eine Kovariante der Funktion f darstellt.

Von Interesse ist auch die Darstellung der In- und Kovarianten durch die Wurzeln $\alpha_1, \alpha_2 \ldots \alpha_n$ der Gleichung $f = 0$, in der man sich f inhomogen geschrieben zu denken hat. Zunächst muß jede Kovariante eine homogene, ganze Funktion der Differenzen $\alpha_i - \alpha_j$ und $\alpha_i - x$ $(i, j = 1, 2, \ldots, n)$ sein. In jedem Gliede dieser Funktion muß ferner sowohl jede der Wurzeln α_i wie auch die Variable x gleich oft vorkommen, außerdem muß der ganze Ausdruck symmetrisch bezüglich der Wurzeln α_i sein, d. h. sich nicht ändern, wenn man die Wurzeln beliebig vertauscht. Sind diese Bedingungen erfüllt, so hat man aber auch immer eine Kovariante oder, wenn die Differenzen $\alpha_i - x$ ganz fehlen, eine Invariante von f vor sich. (Vgl. Salmon, *Modern*

Algebra, Lesson XIII.) Aus irgendeinem Produkte von Wurzeldifferenzen kann man eine Kovariante von f ableiten, indem man derartige Faktoren $x - \alpha_i$ hinzufügt, daß jede Wurzel in dem Ausdrucke eine gleiche Anzahl Male vorkommt, und darauf die Summe dieses Ausdruckes und aller anderen, die aus ihm durch Permutation der Wurzeln entstehen, bildet. Dann ergibt sich immer, falls die Summe nicht identisch verschwindet, eine Kovariante von f. Wenn eine Invariante aus nur einem Gliede besteht, so muß dieses Glied alle Wurzeldifferenzen und zwar jede dieselbe gerade Anzahl Male enthalten. Diese Invarianten sind also alle die Potenzen einer, welche als das Produkt der Quadrate aller Wurzeldifferenzen auftritt und die *Diskriminante* der Gleichung $f = 0$ ist; sie verschwindet, wenn zwei Wurzeln dieser Gleichung einander gleich werden. Sie ist eine Invariante $2(n-1)^{\text{ten}}$ Grades (vgl. S. 274).

Man erkennt leicht, daß der Ausdruck irgendeiner Komitante für die Wurzeldifferenzen mindestens von demselben Grade ist wie die Urform, also vom n^{ten}. Daraus kann man weiter schließen, daß von einer Form n^{ter} Ordnung alle Invarianten verschwinden, wenn sie einen ν-fachen Linearfaktor hat und $\nu > \frac{n}{2}$ ist.

Als eine naturgemäße Erweiterung der einfachen Binärform erscheinen Formen, die *mehrere Paare von Veränderlichen* und zwar jedes Paar in jedem Gliede zu derselben Dimension enthalten, die also aus einer gewöhnlichen Binärform von x_1, x_2 entstehen, indem man in dieser die Koeffizienten wieder als Formen gleicher Ordnung von neuen Variablen y_1, y_2 ansieht usw. Wir wollen eine solche Form eine *Simultanform* nennen.

Man kann nun solche Simultanformen aus einer gewöhnlichen Binärform ableiten durch einen Prozeß, welcher als *Polarenbildung* bezeichnet wird und den wir auch als Plückerschen *Prozeß* bezeichnen können. Er wird in der folgenden Weise definiert:

$$\Delta_y f = \frac{1}{n}\left(\frac{\partial f}{\partial x_1} y_1 + \frac{\partial f}{\partial x_2} y_2\right).$$

Durch r-malige Anwendung resultiert die sogenannte r^{te} Polare der vorgelegten Form n^{ter} Ordnung f:

$$\Delta_y^r f = \frac{1}{n(n-1)\cdots(n-r+1)} \sum_{\varrho=0}^{r} \frac{\partial^r f}{\partial x_1^{r-\varrho}\partial x_2^{\varrho}} y_1^{r-\varrho} y_2^{\varrho}.$$

Nach dieser Definition ist der Satz evident, daß die $(r+s)^{\text{te}}$ Polare die s^{te} Polare der r^{ten} Polare ist. Man erhält die sukzessiven Polaren, indem man die Form $f_\lambda = f(x_1 + \lambda y_1, x_2 + \lambda y_2)$ nach Potenzen von λ entwickelt. Diese Entwicklung kann man schreiben:

$$f_\lambda = f + \binom{n}{1} \Delta_y f \cdot \lambda + \binom{n}{2} \Delta_y^2 f \cdot \lambda^2 + \cdots + \Delta_y^n f \cdot \lambda^n.$$

Der Koeffizient des letzten Gliedes, $\Delta_y^n f$, ist die Form f, nur gebildet für die Variablen y_1, y_2. Bildet man die Form $f(\xi_1 y_1 + \eta_1 y_2, \xi_2 y_1 + \eta_2 y_2)$, d. h. transformiert man die vorgelegte Form durch eine lineare Substitution und schreibt die transformierte Form:

$$f' = a'_0 y_1^n + \binom{n}{1} a'_1 y_1^{n-1} y_2 + \binom{n}{2} a'_2 y_1^{n-2} y_2 + \cdots + a'_n y_2^n,$$

so wird hierin der Koeffizient

$$a'_r = \Delta_\eta^r f(\xi),$$

d. h. die nach η_1, η_2 gebildete r^{te} Polare der für die Variablen ξ_1, ξ_2 genommenen Form f. Die Polaren $U = \Delta_y^r f(x)$ genügen der Differentialgleichung:

$$\frac{\partial^2 U}{\partial x_1 \partial y_2} - \frac{\partial^2 U}{\partial x_2 \partial y_1} = 0.$$

Umgekehrt läßt sich jede Simultanform, welche dieser Differentialgleichung genügt, als Polare einer gewöhnlichen Binärform gewinnen.

Führt man die Polarenbildung an irgendeiner Kovariante einer Simultanform aus (diese Kovariante ist im allgemeinen selbst eine Simultanform), so entstehen neue Kovarianten der Simultanform. *Die Polarenbildung ist ein invarianter Prozeß.*

Die gleiche Eigenschaft kommt der folgenden, an einer Kovariante Q ausgeführten Operation zu:

$$\Omega Q = \frac{1}{pq} \left(\frac{\partial^2 Q}{\partial x_1 \partial y_2} - \frac{\partial^2 Q}{\partial x_2 \partial y_1} \right).$$

Diese Operation heißt der *Caylcysche* oder *Omega-Prozeß* (Cayley, *Cambr. Dubl. Math. Journ.* **1** (1846), 104, *Papers* **1**, 95) p, q sind die Ordnungen der Kovariante für die Variablenpaare x und y. Die charakteristische Eigenschaft einer Polare U drückt sich dann einfach so aus, daß $\Omega U = 0$ wird.

Es ist nun von Wichtigkeit, daß sich jede Simultanform in bestimmter Weise auf Polarformen zurückführen läßt. Diese Zurückführung wird durch die *Clebsch-Gordansche Formel* gegeben. Ist die Simultanform F von der n^{ten} Ordnung für die x und von der m^{ten} für die y $(m < n)$, so wird zunächst die Polare $\Delta_x{}^m F$ eine Form der x allein. Da ferner ΩF von der $(n-1)^{\text{ten}}$ Ordnung für die x und von der $(m-1)^{\text{ten}}$ für die y ist, wird auch $\Delta_x^{m-1}\Omega F$ eine Form der x allein, usw. Die Formen der x, die man so erhält:

$$E_0 = \Delta_x{}^m F,\ E_1 = \Delta_x^{m-1}\Omega F, \ldots E_\varrho = \Delta_x^{m-\varrho}\Omega^\varrho F, \ldots E_m = \Omega^m F,$$

nennt Gordon die *Elementarkovarianten von* F. Dann ergibt sich, wenn man abkürzend

$$x_1 y_2 - x_2 y_1 = (xy)$$

setzt, die in Rede stehende Formel:

$$F = \sum_{\varrho=0}^{m} (m,n)_\varrho \Delta_y^{m-\varrho} E_\varrho \cdot (xy)^\varrho,$$

durch welche die Reduktion der allgemeinen Simultanform F auf Polarformen $\Delta_y^{m-\varrho} E_\varrho$ geleistet wird. Daß noch andere Faktoren, die hier durch die Potenzen des Elementarfaktors (xy) vertreten sind, hinzutreten müssen, ist von vornherein klar, da eine bloße Verbindung von Polarformen wieder eine Polarform wäre. Die Entwicklung der Simultanform F nach Potenzen von (xy), die durch die Clebsch-Gordansche Formel gegeben wird, ist eindeutig bestimmt; die Zahlkoeffizienten sind dabei die folgenden:

$$(m,n)_0 = 1, \quad (m,n)_\varrho = \frac{\binom{m}{\varrho}\binom{n}{\varrho}}{\binom{m+n-\varrho+1}{\varrho}} \qquad (\varrho = 1, 2, \ldots, m).$$

An die Komitanten einer Binärform reihen sich naturgemäß die *simultanen Komitanten mehrerer Binärformen*, die wir auch in Betracht zu ziehen haben.

Wir greifen zunächst besondere Bildungen unter den Komitanten zweier Binärformen f, g von gleich hoher Ordnung heraus. Zu solchen Komitanten kann man gelangen, indem man die Form

$$f_\lambda = f + \lambda g$$

bildet. Ist dann Q eine Komitante k^{ten} Grades von f, so wird

die entsprechende Komitante $Q_{(\lambda)}$ für f_λ von folgender Form:

$$Q_{(\lambda)} = Q + \frac{\lambda}{1!} Q_1 + \frac{\lambda^2}{2!} Q_2 + \cdots \frac{\lambda^k}{k!} Q_k.$$

In diesem Ausdrucke sind $Q_1, Q_2 \ldots$ simultane Komitanten von f und g. Nehmen wir für die Form g die folgende:

$$g = b_0 x_1^n + \binom{n}{1} b_1 x_1^{n-1} x_2 + \binom{n}{2} b_2 x_1^{n-2} x_2^2 + \cdots + b_n x_2^n,$$

so wird

$$Q_{\varrho+1} = \delta Q_\varrho,$$

wenn wir durch das δ-Zeichen den folgenden Prozeß ausdrücken:

$$\delta Q_\varrho = \frac{\partial Q_\varrho}{\partial a_0} b_0 + \frac{\partial Q_\varrho}{\partial a_1} b_1 + \cdots + \frac{\partial Q_\varrho}{\partial a_n} b_n.$$

Diese Operation wird als *Aronholdscher Prozeß* bezeichnet.

Ausgerechnet ergibt sich für Q_ϱ der Wert:

$$Q_\varrho = \sum_{\varkappa, \lambda \cdots \nu} \frac{\partial^\varrho Q}{\partial a_\varkappa \partial a_\lambda \cdots \partial a_\nu} b_\varkappa b_\lambda \cdots b_\nu.$$

Alle aus einer Komitante Q so folgenden Komitanten Q_ϱ heißen nach Cayley deren *Emananten* (Cayley, *Papers* 2, p. 321).

So geht aus jeder Invariante einer Binärform eine Kovariante derselben hervor und jede Kovariante in eine andere über, deren Grad um 1 geringer und deren Ordnung um n höher ist.

Man kann den Aronholdschen Prozeß auch auf irgendeine simultane Komitante Ω zweier Formen gleicher Ordnung f, g anwenden. Hierbei kann insbesondere der Fall eintreten, daß

$$\delta \Omega \equiv 0$$

wird. Dann heißt Ω eine *Kombinante* der beiden Formen f, g. (Vgl. Sylvester, *Cambr. Dubl. M. J.* 8 (1853), 256.) Sie hat die charakteristische Eigenschaft, daß sie sich nur um einen konstanten, d. h. von den Koeffizienten der Formen unabhängigen Faktor ändert, wenn man für f und g irgend zwei Formen aus dem „Formenbüschel" $f + \lambda g$ nimmt. Daraus folgt, daß, wenn wir den früheren Ausdruck für die Komitanten $Q_{(\lambda)}$ der Form $f + \lambda g$ als Funktion von λ betrachten, jede Invariante dieser Funktion eine Kombinante der Formen f, g ist.

Auch ergibt sich, daß jede simultane Komitante der Formen f, g, welche deren Koeffizienten $a_0, a_1, a_2 \ldots$, $b_0, b_1, b_2 \ldots$ nur in den Verbindungen $a_0b_1 - a_1b_0$, $a_0b_2 - a_2b_0$, $a_1b_2 - a_2b_1 \ldots$ enthält, also nur von den Determinanten der Matrix

$$\left\| \begin{matrix} a_0 & a_1 & a_2 & \ldots & a_n \\ b_0 & b_1 & b_2 & \ldots & b_n \end{matrix} \right\|$$

abhängt, eine Kombinante ist.

Gordan, *Math. Ann.* **3** (1871), 355, gewinnt alle Kombinanten aus einer *Fundamentalkombinante*:

$$F = \frac{f(x_1, x_2)\, g(y_1, y_2) - g(x_1, x_2)\, f(y_1, y_2)}{x_1 y_2 - x_2 y_1},$$

deren Invarianten- und Kombinanteneigenschaft offensichtlich ist, indem er diese Simultanform F nach Potenzen von (xy) entwickelt. Die Koeffizienten in dieser Entwicklung sind wieder Kombinanten, aus denen sich alle anderen rational ableiten lassen.

Schreibt man die Simultanform F nach Ausführung der Division in der Gestalt

$$F = \sum_{\varrho, \sigma = 0}^{n-1} c_{\varrho, \sigma} x_1^{\varrho} x_2^{n-1-\varrho} y_1^{\sigma} y_2^{n-1-\sigma},$$

wobei $c_{\varrho\sigma} = c_{\sigma\varrho}$ wird, so ist der Fall besonders merkwürdig, wo die Gleichungen $f(x_1, x_2) = 0$, $g(x_1, x_2) = 0$ eine gemeinsame Wurzel $x_1 : x_2$ haben. Dann muß F für diesen Wert von $x_1 : x_2$ unabhängig von den Werten der y verschwinden. Es bestehen also die n Gleichungen:

$$\sum_{\varrho = 0}^{n-1} c_{\varrho\sigma} x_1^{\varrho} x_2^{n-1-\varrho} = 0 \qquad (\sigma = 0, 1 \ldots n-1)$$

für die n Größen $x_1^{\varrho} x_2^{n-1-\varrho}$. Durch die Elimination dieser n Größen aus den n Gleichungen ergibt sich, daß die Determinante

$$R = |c_{\varrho\sigma}|$$

verschwinden muß. Dies ist die Cayleysche *Form* der *Resultante* zweier Binärformen gleicher Ordnung (Cayley, *Journ. f. Math.* **53** (1857), 366, *Papers* **4**, 38). (Vgl. S. 271.)

Führt man den Aronholdschen Prozeß für die Form f und statt der Form g für die n^{te} Potenz einer Linearform aus,

für welche wir $z_2 x_1 - z_1 x_2$ nehmen, so wird zu setzen sein:

$$(b_0, b_1, b_2 \ldots b_n) = (z_2^n, -z_1 z_2^{n-1}, z_1^2 z_2^{n-2} \ldots (-1)^n z_1^n)$$

und der Aronholdsche Prozeß geht in einen anderen über, der nach Cayley (*Phil. Trans.* **146** (1856), *Papers* **2**, 321) als *Evektantenbildung* bezeichnet wird. Wir wollen ihn durch das Symbol $\mathfrak{E}$ kennzeichnen, so daß die Evektante einer Komitante Q zu setzen ist:

$$\mathfrak{E}Q = \frac{\partial Q}{\partial a_n} z_1^n - \frac{\partial Q}{\partial a_{n-1}} z_1^{n-1} z_2 + \cdots \pm \frac{\partial Q}{\partial a_0} z_2^n .$$

Lassen wir z_1, z_2 mit x_1, x_2 zusammenfallen, so ist dies eine neue Komitante der Grundform f. Der Grad derselben ist um 1 niedriger und ihre Ordnung um n höher als die von Q.

§ 3. Die Cayley-Aronholdsche Symbolik.

Cayley hat in der Arbeit, welche die Grundlage der ganzen Invariantentheorie gebildet hat, auf den nach ihm benannten Ω-Prozeß eine Ableitung der Komitanten gegründet (*Cambr. Dubl. Math. Journ.* **1** (1846), 104, *Journ. f. Math.* **30**, 1, *Papers* **1**, 117), die von den bisher angeführten Methoden zur Aufstellung von In- und Kovarianten wesentlich verschieden ist und in ihrer formellen Weiterbildung durch Aronhold und Clebsch große Bedeutung gewonnen hat. Man vgl. die grundlegende Arbeit von Clebsch, *Journ. f. Math.* **59** (1861), 1, aus welcher die Grundgedanken und das Charakteristische der Methode klarer zu erkennen sind als aus den auf der formalen Vollendung der Theorie basierten Lehrbüchern. Diese Arbeit bildet den Wendepunkt in Clebschs wissenschaftlicher Tätigkeit und inauguriert gleichzeitig die eine Zeitlang in Deutschland bestehende Herrschaft der analytisch-geometrischen Disziplinen. Cayley wendet den Ω-Prozeß auf das Produkt $f \cdot g$ zweier Formen n^{ter} und m^{ter} Ordnung an, von denen er die erste mit den Variablen x_1, x_2, die zweite mit den Variablen y_1, y_2 geschrieben denkt. Es ergibt sich dann:

$$\Omega[f(x) \cdot g(y)] = \frac{1}{n\,m}\left[\frac{\partial f(x)}{\partial x_1} \cdot \frac{\partial g(y)}{\partial y_2} - \frac{\partial f(x)}{\partial x_2} \cdot \frac{\partial g(y)}{\partial y_1}\right].$$

Nun kann man hinterher die y mit den x zusammenfallen lassen. Wir wollen dies ausdrücken, indem wir ω statt Ω schreiben, und finden dann:

$$\omega(f\cdot g)=\{\Omega[f(x)\cdot g(y)]\}_{y=x}=\frac{1}{nm}\left[\frac{\partial f(x)}{\partial x_1}\cdot\frac{\partial g(x)}{\partial x_2}-\frac{\partial f(x)}{\partial x_2}\cdot\frac{\partial g(x)}{\partial x_1}\right].$$

Die so gefundene Kovariante ist die *Jacobische* oder *Funktionaldeterminante* der Formen f, g. Bei wiederholter Anwendung dieses Prozesses ergibt sich:

$$\omega^r(f\cdot g)=\frac{(n-r)!\,(m-r)!}{n!\qquad m!}\sum_{\varrho=0}^{r}(-1)^\varrho\binom{r}{\varrho}\frac{\partial^r f}{\partial x_1^{r-\varrho}\partial x_2^{\varrho}}\cdot\frac{\partial^r g}{\partial x_1^{\varrho}\partial x_2^{r-\varrho}}.$$

Man kann hinterher auch g mit f zusammenfallen lassen und findet so z. B.:

$$\omega^2(f\cdot f)=\frac{2}{n^2}\left[\frac{\partial^2 f}{\partial x_1^2}\cdot\frac{\partial^2 f}{\partial x_2^2}-\left(\frac{\partial^2 f}{\partial x_1\partial x_2}\right)^2\right].$$

Dies ist die *Hessesche Kovariante* von f. Sind f und g insbesondere von derselben Ordnung n, so wird $\omega^n(f\cdot g)$ eine bilineare Invariante, nämlich:

$$\omega^n(f\cdot g)=\sum_{\varrho=0}^{n}(-1)^\varrho\binom{r}{\varrho}a_{r-\varrho}b_\varrho,$$

wobei $a_0, a_1, \ldots, b_0, b_1, \ldots$ die Koeffizienten von f und g sind. Läßt man wieder f und g zusammenfallen, so bekommt man einen von 0 verschiedenen Ausdruck nur, wenn die Ordnung n gerade ist, und zwar wird dann

$$\omega^n(f\cdot f)=2\left\{a_n a_0-\binom{n}{1}a_{n-1}a_1+\cdots\pm\frac{1}{2}\binom{n}{\frac{1}{2}n}a_{\frac{1}{2}n}^2\right\}.$$

Wenn die gefundene bilineare Invariante zweier Formen verschwindet, so heißen die beiden Formen *apolar*. Der Begriff stammt von Battaglini, *Napoli Rendic.* **2** (1863), 158, *Atti* **2** (1865), der Ausdruck von Reye, *Journ. f. Math.* **78** (1874), 97, **79** (1875), 159. Battaglini sagt *harmonisch konjugiert*, Rosanes, *Journ. f. Math.* **76** (1873), 312, *konjugiert*. Die Apolaritätstheorie ist ausführlich dargestellt in dem Buche von Fr. Meyer, *Apolarität und rationale Kurven*, Tübingen 1883.

Jede Binärform von ungerader Ordnung ist zu sich selbst apolar. Damit eine Form f von gerader Ordnung es ist, muß die Invariante $\omega^n(f\cdot f)$ verschwinden, die deshalb von Battaglini *Harmonizante* genannt wird.

Sind zwei Formen f, g apolar und die eine lautet in ihre Linearfaktoren zerfällt

$$f = v_1 v_2 \dots v_n,$$

so wird die andere von der Gestalt

$$g = c_1 v_1^n + c_2 v_2^n + \dots + c_n v_n^n$$

(Rosanes, *Journ. f. Math.* **75** (1873) 166).

Zu irgend n Binärformen n^{ter} Ordnung ist mindestens *eine* Binärform n^{ter} Ordnung apolar, und die ersteren lassen sich nach dem vorigen Satze alle durch die in die n^{ten} Potenzen erhobenen Linearfaktoren der apolaren Form linear ausdrücken.

Wenn jetzt mehr als zwei Formen, f, g, h, ... vorliegen, so schreiben wir diese der Reihe nach mit den Variablen x_1, x_2; y_1, y_2; z_1, z_2, ... und können dann den Ω-Prozeß für irgend zwei dieser Variablenpaare auf das Produkt $f \cdot g \cdot h \cdots$ anwenden. Wir müssen, um auszudrücken, welche der zwei Variablenpaare wir wählen, die Bezeichnung dieser Variablen oder auch der Formen, zu denen sie gehören, als Indizes an Ω setzen, also statt einfach Ω nunmehr schreiben Ω_{fg}, Ω_{fh}, ... Wenden wir nun diese verschiedenen Prozesse eine beliebige Anzahl Male hintereinander an, so ergibt sich zunächst, daß das Resultat von der Reihenfolge, die wir hierbei beobachten, völlig unabhängig ist und nur davon abhängt, wie oft jeder Prozeß im ganzen angewendet worden ist. An die Stelle von

$$\Omega_{fg}^r \, \Omega_{fh}^s \dots [f(x) \cdot g(y) \cdot h(z) \cdots]$$

bringen wir dann wieder

$$\omega_{fg}^r \, \omega_{fh}^s \dots [f(x) \cdot g(x) \cdot h(x) \cdots],$$

indem wir die Variablenpaare x, y, z, ... zusammenfallen lassen. Wenn wir nun symbolisch

$$\omega_{fg} = f_1 g_2 - f_2 g_1 \text{ usw.}$$

setzen, so ergibt sich folgende einfache Regel zur Auswertung des vorstehenden Ausdruckes. Man rechne den Ausdruck

$$(f_1 g_2 - f_2 g_1)^r (f_1 h_2 - f_2 h_1)^s \cdots$$

aus, als ob die Symbole f_1, f_2, ... Zahlen bedeuteten. Dann

ersetze man jedes vorkommende Glied wie

$$\alpha f_1^{\nu_1} f_2^{\nu_2} g_1^{\mu_1} g_2^{\mu_2} h_1^{\lambda_1} h_2^{\lambda_2} \cdots$$

durch das entsprechende Produkt von Differentialquotienten:

$$\alpha \frac{(n-\nu)!}{n!} \frac{(m-\mu)!}{m!} \frac{(l-\lambda)!}{l!} \cdots \frac{\partial^\nu f}{\partial x_1^{\nu_1} \partial x_2^{\nu_2}} \cdot \frac{\partial^\mu g}{\partial x_1^{\mu_1} \partial x_2^{\mu_2}} \cdot \frac{\partial^\lambda h}{\partial x_1^{\lambda_1} \partial x_2^{\lambda_2}} \cdots,$$

wenn n, m, l, ... die Ordnungen von f, g, h, ... bedeuten und auch $\nu = \nu_1 + \nu_2$, $\mu = \mu_1 + \mu_2$, $\lambda = \lambda_1 + \lambda_2$, ... von Glied zu Glied nicht wechseln. Hinterher kann man auch noch die Formen f, g, h, ... zusammenfallen lassen und erhält dann jedesmal eine Komitante der einen Form f.

Besonders zu beachten ist der Fall, daß

$$\nu = n, \ \mu = m, \ \lambda = l, \ \ldots$$

ist. Dann reduzieren sich die in dem Ausdruck für die Komitante der Formen f, g, h, ... vorkommenden Differentialquotienten auf Koeffizienten der betr. Form, indem

$$\frac{1}{n!} \frac{\partial^n f}{\partial x_1^{n-\nu_2} \partial x_2^{\nu_2}} = a_{\nu_2} \text{ usw.}$$

wird, und es entsteht eine simultane Invariante der Formen f, g, h, ..., die in den Koeffizienten jeder dieser Formen linear ist. Aus einer solchen simultanen Invariante mehrerer Formen f, g, h, ... kann man sich aber jede Invariante einer einzigen Form f durch Zusammenfallen der verschiedenen Formen hervorgegangen denken. Man kann nämlich umgekehrt aus der letzteren Invariante eine derartige simultane Invariante durch wiederholte Anwendung des Aronholdschen Prozesses gewinnen. Gehen wir nun auf den symbolischen Ausdruck zurück, aus dem wir die simultane Invariante hergeleitet haben, so können wir ihn durch Einführung der abkürzenden Bezeichnung $f_1 g_2 - f_2 g_1 = (fg)$ usw. schreiben

$$(fg)^r (fh)^s \cdots,$$

und als charakteristisches Merkmal ergibt sich, daß, wenn wir $f \cdot g$ für (fg) setzen usw., herauskommen muß

$$f^n g^n h^n \cdots.$$

Derart ist ein symbolischer Ausdruck für die Invarianten einer Binärform gefunden. Daß in der Tat jede Invariante so darstellbar ist, ergibt sich daraus, daß jede simultane Invariante von $f, g, h, \ldots$ die für die Koeffizienten aller dieser Formen linear ist, wenn man

$$a_\varrho = f_1^{n-\varrho} f_2^\varrho, \; b_\varrho = g_1^{n-\varrho} g_2^\varrho, \ldots$$

setzt, in eine Invariante der linearen Formen

$$f_x = f_1 x_1 + f_2 x_2, \; g_x = g_1 x_1 + g_2 x_2, \ldots$$

d. h. in eine ganze rationale Funktion der Determinanten $(fg), (fh), \ldots$ übergehen muß.

Läßt man die Formen $f, g, h, \ldots$ alle *bis auf eine* zusammenfallen und nimmt für die übrigbleibende Form die folgende:

$$(z_2 x_1 - z_1 x_2)^k,$$

so erhält man, wenn man hinterher für z_1, z_2 wieder x_1, x_2 schreibt, eine Kovariante von f, und andererseits, da man an f_1, f_2; $g_1, g_2, \ldots$ jetzt $z_2, -z_1$ oder $x_2, -x_1$ anzureihen hat, einen symbolischen Ausdruck von der Form

$$(fg)^r (fh)^s \cdots f_x^\varrho g_x^\sigma \cdots, \qquad (\varrho + \sigma + \cdots = p)$$

und dies ist der *allgemeine* symbolische Ausdruck für eine Kovariante einer Binärform, da sich jede solche auf eine simultane Invariante dieser Binärform und der Potenz einer Linearform zurückführen läßt.

Bei dieser symbolischen Schreibweise der Kovarianten ist die Grundform selbst, die ja mit zu diesen Kovarianten gehört, zu schreiben:

$$f = f_x^n = g_x^n = \cdots$$

Die so gegebene Darstellung der Komitanten hat den Vorzug, deren Bildungsweise aus einem verhältnismäßig einfachen Ausdrucke unmittelbar und vollständig erkennen zu lassen. Wie kompliziert die wirklichen Ausdrücke werden, kann man aus dem von Salmon im Anhange der *Lessons on Algebra* gegebenen Beispiele einer Invariante 15. Grades einer Form sechster Ordnung erkennen, die ausgeschrieben 13 Seiten einnimmt.

Die auseinandergesetzte Methode der Aufstellung von Komitanten läßt sich als *direkte* von einer *indirekten* Methode unter-

scheiden, diese beruht auf einer fortlaufenden Verwendung des ω-Prozesses, der von Clebsch als *Überschiebung* (*Binäre Formen*, S. 99) und später auch als *Clebschscher Prozeß* beezeichnet wurde. Man kann nämlich zuerst bilden $\omega(f \cdot f)$, $\omega^2(f \cdot f), \ldots$, darauf $\omega^s(\omega^r(f \cdot f) \cdot f)$ für $s = 1, 2, \ldots$ und so fortfahren, dann gilt der wichtige Satz, der von Gordan herrührt, *daß man jede Komitante durch sukzessive Anwendung des Überschiebungsprozesses gewinnen kann.*

Die Kovariante $\omega^r(f \cdot g)$, die Clebsch noch einfacher $(f, g)^r$ schreibt, wird in der auseinandergesetzten symbolischen Darstellung

$$(f, g)^r = (fg)^r f_x^{n-r} g_x^{n-r}.$$

Diese Darstellung bleibt äußerlich vollkommen ungeändert, wenn man die Formen f und g zusammenfallen läßt.

Die symbolische Darstellung läßt sich unmittelbar auch auf Formen mit mehreren Paaren von Veränderlichen erweitern. So wird z. B. die Darstellung der r^{ten} Polare der Form f

$$\Delta_y^r f = f_x^{n-r} f_y^r.$$

Eine allgemeine Simultanform F mit zwei Paaren von Veränderlichen ist zu schreiben:

$$F = f_x^n g_y^m,$$

und wendet man auf diese Form den Ω-Prozeß an, so ergibt sich symbolisch:

$$\Omega F = (fg) f_x^{n-1} g_y^{m-1},$$

durch r malige Anwendung:

$$\Omega^r F = (fg)^r f_x^{n-r} g_y^{m-r}.$$

In diesem symbolischen Kalkül wird die allgemeine Simultanform F von dem Produkt zweier Binärformen, die eine, f_x^n, für Variable x, die andere, g_y^m, für Variable y geschrieben, nicht unterschieden.

Eine Simultanform, die für beide Variablenpaare von gleicher Ordnung ist, besitzt zwei ausgezeichnete Bildungen, die von Gordan (*Math. Ann.* **3** (1871), 355) herrühren und zu denen wir auf folgende Art gelangen. Wir nehmen die m Formen:

$$f_x^{(i)m} g_y^{(i)m} \qquad (i = 1, 2, \ldots, m),$$

die wir hinterher zusammenfallen lassen, und bilden die Invariante:

$$R = \prod_{i,j} (f^{(i)} f^{(j)}) (g^{(i)} g^{(j)}) \qquad (i, j = 1, 2, \ldots, m)$$

und die Kovariante:

$$\Theta = \prod_{i,j} (f^{(i)} f^{(j)}) (g^{(i)} g^{(j)}) f_x^{(i)} g_x^{(j)} \qquad (i, j = 1, 2, \ldots, m-1).$$

Wählen wir für die Simultanform die oben besprochene Fundamentalkombinante, wobei $m = n - 1$ wird, so ist R die Resultante der zugrunde liegenden Formen f, g, also $R = 0$ die Bedingung, daß sie einen gemeinsamen Linearfaktor besitzen. Das identische Verschwinden von Θ aber wird die Bedingung, daß f und g *zwei* Linearfaktoren gemein haben. Stephanos (*Ann. de l'École norm.* (3) **1** (1884), 329) hat bewiesen, daß Θ die einzige Form $2(n-1)^{\text{ter}}$ Ordnung ist, deren n^{te} Überschiebungen mit den Grundformen f und g identisch verschwinden. Die Kovariante Θ ist eine Kombinante der Formen f und g.

Die zu $n-1$ Formen n^{ter} Ordnung apolaren Formen lassen sich in der Gestalt $f + \lambda g$ darstellen, sie bilden ein Formenbüschel. Die Kovariante Θ der Formen f, g ist als Kombinante eine Kovariante des ganzen Formenbüschels (d. h. irgend zweier Formen aus ihm) und damit auch den $n-1$ vorgelegten Formen in eindeutiger Weise zugeordnet. Es gilt dann der merkwürdige Satz, daß sich die $n-1$ Formen als lineare Kombinationen der $(n-2)^{\text{ten}}$ Derivierten von Θ nach den Veränderlichen x_1, x_2 darstellen lassen (Berzolari, *Napoli Rendic.* (2) **5** (1891), 35).

Die Regeln, die für den symbolischen Kalkül gelten, können nach dem Auseinandergesetzten keine anderen sein als die Regeln, die bei der wirklichen Zahlenrechnung für die Komitanten einer Reihe von linearen Formen

$$f_x, g_x, h_x, k_x, l_x, \ldots$$

gelten. Aber alle diese Formen werden hinterher als identisch angenommen. Es kann also ein Ausdruck, der symbolische Bedeutung hat, bei Vertauschung der Symbole $f, g, \ldots$ sich nicht ändern, und wechselt er hierbei sein Vorzeichen, so muß er identisch verschwinden. Die symbolischen Rechenregeln lassen sich auf die eine Regel zurückführen:

$$\text{(I)} \qquad f_x g_y - g_x f_y = (fg)(xy),$$

die den Fundamentalsatz für das Operieren mit linearen Binärformen darstellt. Macht man in ihr $y_1 = -h_2$, $y_2 = h_1$, so erhält man:

$$(\text{II}) \qquad f_x(gh) + g_x(hf) + h_x(fg) = 0.$$

Macht man hierin weiter $x_1 = -k_2$, $x_2 = k_1$, so entsteht:

$$(\text{III}) \qquad (fk)(gh) + (gk)(hf) + (hk)(fg) = 0.$$

Die invarianten Prozesse hat Gordan auf einen zurückgeführt, der bei Verwendung dieser Symbolik in formaler Hinsicht der einfachste ist. Er besteht darin, daß in dem symbolischen Ausdrucke $f_x g_x$ durch (fg) ersetzt wird. Diesen Prozeß bezeichnet Gordan als *Faltung*.

Auf die symbolische Darstellung der Komitanten läßt sich der Beweis eines wichtigen Satzes gründen, der als der *Hermitesche Reziprozitätssatz* (Hermite, *Cambr. Dubl. Math. Journ.* **9** (1854), 172, man vgl. Salmon, *Higher Algebra*, Lesson XIV, ferner die Erweiterung dieses Theorems betreffend Deruyts, *Bull. de l'Acad. de Bruxelles* (3) **22** (1891) und Gordan, *Gött. Nachr.* (1897), 182) bezeichnet wird. Wir setzen in dem allgemeinen symbolischen Ausdrucke einer Kovariante der Binärform f

$$-\frac{f_2}{f_1} = \alpha, \quad -\frac{g_2}{g_1} = \beta, \quad -\frac{h_2}{h_1} = \gamma, \quad \cdots \quad \frac{x_1}{x_2} = x,$$

so wird er, von einem Faktor abgesehen:

$$(\alpha - \beta)^r (\alpha - \gamma)^s \ldots (x - \alpha)^\varrho (x - \beta)^\sigma \ldots,$$

d. h. von genau derselben Form wie der oben mit Hülfe der Wurzeln $\alpha, \beta, \ldots$ von $f = 0$ gebildete Ausdruck. Aber hier ist die Anzahl Male n, die α und ebenso $\beta, \gamma, \ldots$ vorkommt, nicht der Grad der Komitante, sondern die Ordnung der Grundform f. Zu einer Komitante k^{ten} Grades p^{ter} Ordnung einer Form n^{ter} Ordnung findet man so immer eine Komitante n^{ten} Grades p^{ter} Ordnung einer Form k^{ter} Ordnung. Einem System linear unabhängiger oder wie man sagt *asyzygetischer* Komitanten gleicher Ordnung p wird auf diese Weise wieder ein System asyzygetischer Komitanten von derselben Ordnung p zugeordnet.

Als Beispiel der Methode, die Komitanten durch sukzessive Überschiebung zu gewinnen, wollen wir das *System zweier kubischen Formen* f, g behandeln.

Durch direkte Anwendung des Überschiebungsprozesses auf die Formen selbst gewinnt man eine Kovariante vierter Ordnung

$$\vartheta = (f, g), \tag{1}$$

drei quadratische Kovarianten

$$A = (f, f)^2,\ B = (g, g)^2,\ C = (f, g)^2 \tag{2}$$

und eine Invariante

$$J = (f, g)^3. \tag{3}$$

Aus den drei quadratischen Formen A, B, C leitet man durch doppelte Überschiebung derselben über sich selbst und übereinander die sechs Invarianten her:

$$\begin{aligned} D_{11} &= (A, A)^2,\ D_{22} = (B, B)^2,\ D_{33} = (C, C)^2, \\ D_{23} &= (B, C)^2,\ D_{31} = (C, A)^2,\ D_{12} = (A, B)^2, \end{aligned} \tag{4}$$

die mit den früher so bezeichneten Invarianten dreier quadratischer Formen übereinstimmen. Die einmalige Überschiebung der Formen A, B gibt eine quadratische Kovariante:

$$E = (A, B). \tag{5}$$

Überschiebung der quadratischen Formen A, B mit den Urformen führt zu vier Kovarianten dritter Ordnung:

$$\begin{aligned} P &= (f, A),\ Q = (g, B), \\ U &= (f, B),\ V = (g, A), \end{aligned} \tag{6}$$

und doppelte Überschiebung zu den linearen Kovarianten:

$$p = (f, A)^2,\ q = (g, B)^2. \tag{7}$$

Gleichzeitig wird:

$$(f, C)^2 = -\frac{1}{2}p,\ (g, C)^2 = -\frac{1}{2}q.$$

Die Überschiebung der Kovarianten p, q übereinander liefert die Invariante:

$$\Omega = (p, q). \tag{8}$$

Die Überschiebung der linearen Kovarianten p, q mit den

quadratischen Formen A, B, C liefert neue lineare Kovarianten:

$$(9) \qquad p' = (A, p),\ q' = (A, q),\ p'' = (B, p),\ q'' = (B, q).$$

Endlich ergibt die doppelte Überschiebung einer der Urformen f, g mit der kubischen Kovariante P oder Q der anderen die quadratischen Kovarianten:

$$(10) \qquad A' = (P, g)^2,\ B' = (Q, f)^2.$$

Zwischen den gefundenen acht Invarianten J, D_{ij}, Ω bestehen die Relationen:

$$\Omega^2 = \frac{1}{2}\begin{vmatrix} D_{11} & D_{12} & D_{13} \\ D_{21} & D_{22} & D_{23} \\ D_{31} & D_{32} & D_{33} \end{vmatrix},$$

$$2J\Omega = (D_{11}D_{22} - D_{12}^2) - 4(D_{13}D_{32} - D_{33}D_{12}),$$
$$J^2 = 2(D_{12} - D_{33}),$$

und die Resultante der Formen f, g wird:

$$R = 27\Omega - 4J^3.$$

Aus den zwei kubischen Formen f, g kann man die Kovariantenform vierter Ordnung ableiten

$$\mathfrak{f} = fq + gp.$$

Deren Invarianten werden

$$i_{\mathfrak{f}} = -2\Omega J,\quad j_{\mathfrak{f}} = 2\Omega^2,$$

und ihre Kovarianten

$$H_{\mathfrak{f}} = -4\Omega\vartheta,$$
$$T_{\mathfrak{f}} = 2\Omega\{g(Ap + Cq) - f(Bp + Cq)\}.$$

Das System zweier kubischen Formen hat Clebsch behandelt (*Journ. f. Math.* **67** (1867), 360) und v. Gall (*Math. Ann.* **31** (1888), 424); man vgl. außerdem Sylvester, *C. R.* **89** (1879), 828, d'Ovidio, *Torino Atti* **15** (1880), 267 und Gordans *Vorlesungen* **2**, § 32f. Das System dreier kubischen Formen hat v. Gall untersucht (*Math. Ann.* **45** (1894), 207) und das System beliebig vieler kubischen Formen Peano (*Torino Atti* **17** (1882), 580).

Das System zweier Formen vierter Ordnung wurde untersucht von Gordan, *Math. Ann.* **2** (1870), 227, d'Ovidio, *Torino Atti* **15** (1880), 301, 385, 471, **28** (1893), 447, Stroh, *Math. Ann.* **22** (1883), 290, v. Gall, ib. **33** (1889), 197, **34** (1889), 332, Brioschi, *Torino Atti* **31** (1896), 441.

Über Formen mit gemeinsamen Linearfaktoren sehe man Gordan, *Math. Ann.* **3**, Pascal, *Ann. di mat.* (2) **16** (1888), 85, *Napoli Rendic.* (2) **2** (1888), 402, *Lomb. Ist. Rendic.* **37** (1904), 917, 980.

Das System einer Form *fünfter* Ordnung findet man bei Clebsch, *Binäre Formen,* § 73—75. Vgl. auch d'Ovidio, *Torino Atti* **15** (1880), 591, Hammond, *Amer. Journ. of Math.* **8** (1886), 19, *Lond. Math. Soc. Proc.* **27** (1896), 392, Stroh, *Math. Ann.* **33** (1889), 61, **34** (1889), 354. Die Diskriminante berechnete bereits Salmon (*Cambr. Dubl. Math. Journ.* **5** (1850), 159). Die Diskriminante der Form sechster Ordnung hat Brioschi gefunden (*Ann. di mat.* (2) **1** (1868), 159, *Opere* **2**, 77), die Diskriminante einer Form siebenter Ordnung haben Gordan, *Math. Ann.* **31** (1888), 566, und Brioschi, *Ann. di mat.* (2) **26** (1898), 255, untersucht.

Das System einer Form *sechster* Ordnung findet man behandelt bei Clebsch, *Binäre Formen,* § 76ff. Über das System einer Form *siebenter* Ordnung kann man sich orientieren bei v. Gall, *Math. Ann.* **31** (1888), 318.

Das System einer Form *achter* Ordnung hat v. Gall untersucht (*Math. Ann.* **17** (1881), 31, 139).

Neben der Clebsch-Gordanschen Richtung, welche die unabhängigen Komitanten gegebener Formen in ihrem symbolischen Ausdrucke wirklich hinzuschreiben strebt, steht die abzählende Richtung von Cayley und Sylvester, welche nur die Anzahl und den Typus der unabhängigen Komitanten zu bestimmen sucht. Diese Bestimmung beruht auf der Verwendung einer „erzeugenden Funktion", d. h. einer rationalen Funktion von zwei Variablen x, a, in deren Entwicklung der Koeffizient von $a^k x^h$ die Anzahl der linear unabhängigen (asyzygetischen) Komitanten vom Grade k und der Ordnung h (deg-order k, h) liefert. Durch Siebung (tamisage) der so gefundenen Zahlen kann man zu den Anzahlen der von allen niedrigeren unabhängigen Komitanten gelangen. So gelang es Sylvester, Tabellen für die charakteristischen Zahlen der unabhängigen Komitanten aufzustellen, und er hat im *Am. Journ. of Math.* **2** (1879), 223 diese Tabellen für die Binärformen erster bis neunter Ordnungen gegeben.

§ 4. Kanonische und typische Darstellung. Erweiterungen des Invariantenbegriffes.

Eine besondere Aufgabe der Formentheorie besteht darin, die vorgelegte Grundform selbst so umzugestalten, daß sie aus einfacheren Formen in einer besonders charakteristischen Weise zusammengesetzt erscheint. Insbesondere kann sie als Summe von Potenzen linearer Formen dargestellt werden, und dies ist es, was man im engsten Sinne unter der *kanonischen Darstellung* einer Binärform versteht. Die Theorie dieser Darstellungsarten schuf Sylvester, *Cambr. Dubl. Math. Journ.* **6** (1851), 186, *Philos. Magaz.* (4) **2** (1851), 391, auch in einer selbständigen Schrift *An Essay on Canonical Forms* 1851, alles vereinigt in den *Werken* **1** (1904). Die Darstellung einer kubischen Form als Summe dritter Potenzen hatte schon Hesse erörtert (*Journ. f. Math.* **38** (1849), 262, *Werke,* 217). Bei der Ausführung dieser kanonischen Darstellung ist zu scheiden zwischen Formen ungerader und gerader Ordnung.

Jede Form f von der *ungeraden* Ordnung $n = 2\nu + 1$ läßt sich immer als Aggregat von $\nu + 1$ Potenzen linearer Formen folgendermaßen ausdrücken:

$$f = \sum_{\varrho=0}^{\nu} b_\varrho (x_1 - \alpha_\varrho x_2)^{2\nu+1}.$$

Hierbei sind die Koeffizienten α_ϱ die Wurzeln einer Gleichung $(\nu+1)^{\text{ten}}$ Grades

$$A_0 - A_1\alpha + A_2\alpha^2 - \cdots + (-1)^{\nu+1} A_{\nu+1}\alpha^{\nu+1} = 0,$$

der *Sylvesterschen Kanonisante.* Die Symbole $A_0, A_1, \ldots, A_{\nu+1}$ bedeuten die Determinanten der Matrix:

$$\begin{Vmatrix} a_0 & a_1 & \cdots & a_{\nu+1} \\ a_1 & a_2 & \cdots & a_{\nu+2} \\ \cdots & \cdots & \cdots & \cdots \\ a_\nu & a_{\nu+1} & \cdots & a_{2\nu+1} \end{Vmatrix}.$$

Wenn eine Form f von der geraden Ordnung $n = 2\nu$ als Aggregat von ν Potenzen darstellbar sein soll, so muß die Determinante

$$\begin{vmatrix} a_0 & a_1 & \dots & a_\nu \\ a_1 & a_2 & \dots & a_{\nu+1} \\ \dots & \dots & \dots & \dots \\ a_\nu & a_{\nu+1} & \dots & a_{2\nu} \end{vmatrix},$$

die *Katalektikante,* verschwinden. Ist dieselbe nicht Null, so ersetze man, wenn $a_{2\nu} \neq 0$ ist, die Form f durch $f - \lambda x_2^{2\nu}$ und bestimme λ so, daß die Katalektikante für diese neue Form verschwindet. Dann läßt sich $f - \lambda x_2^n$ als Summe von ν Potenzen darstellen, f mithin als Aggregat von $\nu + 1$ Potenzen, und dies ist immer auf unendlich viele Arten möglich, denn wenn man f linear transformiert, so kann man auf unendlich viele Arten erreichen, daß der Koeffizient von a_0 nicht verschwindet, weil hierzu nur erforderlich ist, daß $x_2 = 0$ *keine* Wurzel der Gleichung $f = 0$ ist.

Der kanonischen Darstellung einer Binärform steht die *typische Darstellung* zur Seite. Hierunter versteht man eine solche Umgestaltung der gegebenen Form f, daß die Veränderlichen Kovarianten und die Koeffizienten Invarianten derselben werden. Diese Darstellung findet man ausführlich abgehandelt bei Clebsch, *Binäre Formen,* 7.—9. Abschnitt; vgl. auch Gordans *Vorlesungen.*

Bei einer Form von *ungerader* Ordnung n führt man die typische Darstellung mit Hülfe zweier (für $n > 3$ immer existierenden) *linearen* Kovarianten α_x, β_x aus. Es gilt, wenn f_x irgendeine dritte lineare Form ist, die Identität:

$$f_x(\alpha\beta) = \alpha_x(f\beta) - \beta_x(f\alpha).$$

Erhebt man diese in die n^{te} Potenz, so erhält man:

$$f_x^n(\alpha\beta)^n = \sum_{\varrho=0}^{n} (-1)^\varrho \binom{n}{\varrho} (f\beta)^{n-\varrho} (f\alpha)^\varrho \alpha_x^{n-\varrho} \beta_x^\varrho.$$

Diese Gleichung kann man symbolisch deuten, indem man

$$f = f_x^n$$

annimmt, und erhält dann mit Hülfe der Invarianten

$$C_\varrho = (-1)^\varrho (f\beta)^{n-\varrho} (f\alpha)^\varrho$$

die typische Darstellung:

$$(\alpha\beta)^n \cdot f = \sum \binom{n}{\varrho} C_\varrho \alpha_x^{n-\varrho} \beta_x^\varrho.$$

Bei einer Form von der *geraden* Ordnung $n = 2\nu$ existieren für $n > 4$ immer drei *quadratische* Kovarianten, die wir schreiben:

$$\xi = \alpha_0 x_1^2 + 2\alpha_1 x_1 x_2 + \alpha_2 x_2^2,$$
$$\eta = \beta_0 x_1^2 + 2\beta_1 x_1 x_2 + \beta_2 x_2^2,$$
$$\zeta = \gamma_0 x_1^2 + 2\gamma_1 x_1 x_2 + \gamma_2 x_2^2.$$

Ist R die aus den Koeffizienten $\alpha_0, \alpha_1, \alpha_2, \ldots$ dieser Gleichungen gebildete Determinante, so wird für irgendeine lineare Form $f_x = f_1 x_1 + f_2 x_2$:

$$\begin{aligned} R f_x^2 = {} & [(\beta_1\gamma_2 - \beta_2\gamma_1) f_1^2 + (\beta_2\gamma_0 - \beta_0\gamma_2) f_1 f_2 + (\beta_0\gamma_1 - \beta_1\gamma_0) f_2^2]\,\xi \\ & + [(\gamma_1\alpha_2 - \gamma_2\alpha_1) f_1^2 + (\gamma_2\alpha_0 - \gamma_0\alpha_2) f_1 f_2 + (\gamma_0\alpha_1 - \gamma_1\alpha_0) f_2^2]\,\eta \\ & + [(\alpha_1\beta_2 - \alpha_2\beta_1) f_1^2 + (\alpha_2\beta_0 - \alpha_0\beta_2) f_1 f_2 + (\alpha_0\beta_1 - \alpha_1\beta_0) f_2^2]\,\zeta. \end{aligned}$$

Erhebt man diese Gleichung in die ν^{te} Potenz, so geht sie über in eine Gleichung von der Form:

$$R^\nu f_x^n = \sum_{\varrho,\sigma} C_{\varrho\sigma} \xi^\varrho \eta^\sigma \zeta^{\nu-\varrho-\sigma}.$$

Deutet man diese Formel wieder symbolisch, ersetzt also in den $C_{\varrho\sigma}$ die $f_1^{n-\varrho} f_2^\varrho$ durch a_ϱ, so bedeuten die $C_{\varrho\sigma}$ Invarianten der Form $f = f_x^n = \sum \binom{n}{\varrho} a_\varrho x_1^{n-\varrho} x_2^\varrho$, und man hat eine typische Darstellung vor sich. Zwischen ξ, η, ζ besteht wie zwischen irgend drei quadratischen Formen eine Identität von der Gestalt:

$$c_{11}\xi^2 + c_{22}\eta^2 + c_{33}\zeta^2 + 2c_{23}\eta\zeta + 2c_{31}\zeta\xi + 2c_{12}\xi\eta = 0.$$

Eine andere Art der typischen Darstellung rührt von Hermite her. Sie steht in engem Zusammenhange mit der Theorie der Seminvarianten. Wir können unter den letzteren die folgenden Bildungen als Sylvestersche „Grundformen" auszeichnen:

$$\mathfrak{G}_{2\nu} = \sum_{k=0}^{2\nu} (-1)^k \binom{2\nu}{k} a_k a_{2\nu-k},$$

$$\mathfrak{H}_{2\nu+1} = \sum_{k=0}^{2\nu} (-1)^k \binom{2\nu}{k} a_k (a_0 a_{2\nu+1-k} - a_1 a_{2\nu-k}).$$

Die $\mathfrak{G}_{2\nu}$ sind vom zweiten, die $\mathfrak{J}_{2\nu+1}$ vom dritten Grade. Führen wir noch ein den Ausdruck:

$$\mathfrak{G}_{2\nu+1} = \sum_{k=0}^{2\nu} (-1)^k \binom{2\nu}{k} a_k a_{2\nu+1-k},$$

so können wir einfach setzen:

$$\mathfrak{J}_{2\nu+1} = a_0 \mathfrak{G}_{2\nu+1} - a_1 \mathfrak{G}_{2\nu}.$$

Hermite schlägt nun (*Journ. f. Math.* **52** (1856), 18) folgenden Weg zur Transformation einer Binärform n^{ter} Ordnung f ein. Er setzt für x_1, x_2 ein:

$$x_1 X - \frac{1}{n}\frac{\partial f}{\partial x_2} Y, \quad x_2 X + \frac{1}{n}\frac{\partial f}{\partial x_1} Y$$

und entwickelt den entstehenden Ausdruck nach Potenzen von X, Y, indem er dem Resultat dieser Entwicklung die Form gibt:

$$f X^n + \binom{n}{1} f' X^{n-1} Y + \binom{n}{2} f'' X^{n-2} Y^2 + \cdots.$$

Er bezeichnet hierin $f', f'', \ldots$ als *assoziierte* Kovarianten. Dieselben sind, wie sich herausstellt, alle durch f teilbar und f' verschwindet identisch.

Man kann aber diese Darstellung so wenden, daß man erst in der Urform f die Variablen y_1, y_2 schreibt und sodann:

$$f(x) \cdot y_1 = x_1 \xi - \frac{1}{n}\frac{\partial f}{\partial x_2} \eta, \quad f(x) \cdot y_2 = x_2 \xi + \frac{1}{n}\frac{\partial f}{\partial x_1} \eta$$

setzt, woraus:

$$\xi = \frac{\partial f}{\partial x_1} y_1 + \frac{\partial f}{\partial x_2} y_2, \quad \eta = x_1 y_2 - x_2 y_1$$

folgt, so daß ξ, η die Invarianteneigenschaft haben. Nun ergibt sich:

$$f(x)^{n-1} f(y) = G_0 \xi^n + \binom{n}{1} G_1 \xi^{n-1}\eta + \binom{n}{2} G_2 \xi^{n-2}\eta + \cdots + G_n \eta^n.$$

Hierbei sind die G_ϱ Kovarianten, und insbesondere wird $G_0 = 1$, $G_1 = 0$. Jede ganze rationale Funktion der G_ϱ ist eine Komitante der Urform f, wobei sich eine Potenz von $f(x)$ ausscheidet,

und umgekehrt muß jede Komitante von f sich als eine rationale Funktion der G_ϱ darstellen lassen, in deren Nenner eine Potenz von f steht. Die Formen G nennt Gordan die *Schwesterformen* von f. In symbolischer Darstellung wird, wenn wir die Formen

$$f_x^n,\ f_x'^n,\ f_x''^n,\ \ldots$$

alle mit f zusammenfallen lassen:

$$f \cdot G_\varrho = (ff')(ff'') \ldots (ff^{(k)}) f_x'^{n-1} \ldots f_x^{(k)n-1} f_x^{n-k}.$$

Es werden nun die Fundamentalkovarianten $J_{2\nu}$, $J_{2\nu+1}$ aufgestellt, die durch die Gleichungen definiert sind:

$$f^{2\nu-2} J_{2\nu} = \sum_{k=0}^{2\nu} (-1)^k \binom{2\nu}{k} G_k G_{2\nu-k},$$

$$f^{2\nu-2} J_{2\nu+1} = \sum_{k=0}^{2\nu} (-1)^k \binom{2\nu}{k} G_k G_{2\nu+1-k}.$$

Durch Überschiebungen erhält man die Kovarianten J wie folgt:

$$J_{2\nu} = (f, f)^{2\nu},\ J_{2\nu+1} = (J_{2\nu}, f).$$

Insbesondere wird J_2 gleich der Hesseschen Kovariante von f. Wenn wir die oben eingeführten Seminvarianten für die transformierte Form bilden, so werden die dort mit $\mathfrak{G}_{2\nu}$ bezeichneten unmittelbar $= f^{2\nu-2} J_{2\nu}$, aber mit Rücksicht darauf, daß hier an Stelle von a_0, a_1 $G_0 = 1$, $G_1 = 0$ tritt, werden auch die Seminvarianten $\mathfrak{J}_{2\nu+1} = f^{2\nu-2} J_{2\nu+1}$. Die fundamentalen Kovarianten der Urform fallen somit bis auf den hinzutretenden Faktor $f^{2\nu-2}$ mit den fundamentalen Seminvarianten der transformierten Form zusammen.

Jede isobare Funktion der Koeffizienten einer Binärform hat die Eigenschaft, daß sie sich bei Ausführung der linearen Transformation

$$x_1 = x_1',\ x_2 = \delta x_2'$$

nur um einen konstanten Faktor, nämlich δ^w, wenn w das Gewicht bezeichnet, ändert.

Die Seminvarianten genügen der weitergehenden Forderung, daß sie bei allen linearen Transformationen

$$x_1 = x_1' + \beta x_2',\ x_2 = \delta x_2'$$

sich nur um den Faktor δ^w ändern. Diese Bildungen können

demnach sozusagen als Vorstufen der invarianten Bildungen angesehen werden.

Man kann den Invariantenbegriff so verallgemeinern, daß man an die Stelle der Gesamtheit *aller* linearen Transformationen eine *Gruppe* von solchen Transformationen setzt und die Forderung der Invarianz nur für diese Gruppe aufstellt. Man redet dann von *Invarianten der betr. Gruppe.* Besonders wichtig ist der Fall, wo die Gruppe durch die Forderung festgelegt wird, daß eine bestimmte quadratische Form für alle zu der Gruppe gehörenden Transformationen die Eigenschaft der Invarianz haben soll. Wird für diese *identische Kovariante*, welche die Gruppe festlegt, insbesondere die Quadratsumme $x_1^2 + x_2^2$ gewählt, so spricht man von *orthogonalen Transformationen*, die hier im binären Gebiet von der Form

$$x_1 = \alpha x_1' + \beta x_2', \quad x_2 = -\beta x_1' + \alpha x_2'$$

sind, und entsprechend von orthogonalen Invarianten. Z. B. besitzt eine quadratische Form $a_0 x_1^2 + 2a_1 x_1 x_2 + a_2 x_2^2$ außer ihrer Diskriminante die orthogonale Invariante $a_0 + a_2$. Man vgl. Lie-Scheffers, *Vorlesungen über kontinuierliche Gruppen,* Leipzig 1893, Kap. 23.

Der Begriff der Invarianten kann ferner in der Weise verallgemeinert werden, daß man außer *rationalen* Invarianten auch *algebraische* Invarianten in Betracht zieht, die durch eine algebraische Gleichung mit rationalen Invarianten als Koeffizienten festgelegt werden.

Durch die von Sylvester eingeführten *Reziprokanten* wird der Invariantenbegriff auf irgendwelche fortgesetzt differenzierbare Funktionen y eine Variablen x erweitert. Es handelt sich hier um solche rationale Funktionen der sukzessiven Ableitungen

$$y' = \frac{dy}{dx}, \quad y'' = \frac{d^2y}{dx^2}, \cdots,$$

welche sich nur um einen Faktor, nämlich eine positiv oder negativ genommene Potenz von y', ändern, wenn man die Ableitungen $y', y'' \ldots$ durch die umgekehrten:

$$x' = \frac{dx}{dy}, \quad x'' = \frac{d^2x}{dy^2}, \cdots$$

ersetzt. Die am längsten bekannte Reziprokante ist die Schwarzsche Form:

$$2y'y''' - 3y''^2 = -y'^6(2x'x''' - 3x''^2).$$

Ebenso wird z. B.

$$3y''y^{IV} - 5y'''^2 = y'^8(3x''x^{IV} - 5x'''^2).$$

Diese Reziprokanten sind homogene ganze Funktionen der Ableitungen. Von solchen Reziprokanten gilt, daß sie isobar sind, wenn man als das Gewicht der Derivierten $y^{(m)}$ die Zahl $m - 2$ ansieht. So wird in den vorstehenden Beispielen, bei denen der *Grad* $k = 2$ ist, das Gewicht w das erste Mal $= 0$, das zweite Mal $= 2$. Der Exponent μ der Potenz y'^{μ}, mit der sich die Reziprokante multipliziert, wird allgemein

$$\mu = 3k + w.$$

Aus der relativen Reziprokante wird eine *absolute Reziprokante* R abgeleitet, indem man sie durch $y'^{\frac{1}{2}\mu}$ dividiert. Dann ist $\frac{dR}{dx}$ wieder eine Reziprokante. Die Reziprokanten ändern sich nicht, wenn:

$$x = x_1 + a, \quad y = y_1 + b$$

gesetzt wird (*Translations-Reziprokanten*). Eine Reziprokante, die y' nicht enthält, heißt eine *reine* Reziprokante. Sie bleibt invariant für die Transformationen:

$$x = ax_1 + by_1 + c, \quad y = dx_1 + ey_1 + f$$

(*Affinitäts-Reziprokante*). In diesem Sinne sind die Reziprokanten in der Tat den Invarianten analog.

Vgl. die von Hammond bearbeiteten Vorlesungen Sylvesters *Am. Journ. of Math.* **8** (1886), 196 und ebenda **9** (1887), 1, 113, 297, **10** (1887), 1, ferner Cayley, *Quarterly Journ.* **26** (1893), 169, 289, *Papers* **13**, 366.

§ 5. Formenmoduln.

Die Binärformen finden ihre einfachste geometrische Interpretation durch Punktsysteme auf rationalen Kurven. Von diesem Standpunkte aus liegt es nahe, als natürliche Erweiterung der Binärformen die algebraischen Gebilde anzusehen, deren geometrisches Äquivalent Punktsysteme auf einer beliebigen algebraischen Kurve sind. Nachdem die Theorie der Punktsysteme erst auf transzendentem Wege (Clebsch und Gordan, *Theorie der Abelschen Funktionen*, 1866) in Angriff genommen und darauf die

algebraische Theorie durch Brill und Noether (vgl. deren Bericht über die Entwicklung dieser Theorie, *Math. Ver.* **3**, 1894) begründet worden, haben Kronecker einerseits (*Journ. f. Math.* **92** (1882), 1, *Werke* **2** (1897), 237) und Dedekind-Weber andererseits (*Journ. f. Math.* **92** (1882), 181) eine strenge arithmetische Theorie dieser Gebilde gegeben.

Der allgemeine Begriff, auf dem die letzteren Theorien fußen, ist bei Kronecker das Modulsytem, bei Dedekind-Weber der Modul, und diese beiden Begriffe stehen in engster Beziehung zueinander. In die eigentliche Formentheorie hat Hilbert (*Math. Ann.* **36** (1890), 473) die Moduln eingeführt, die wir zur unzweideutigen Charakterisierung hier *Formenmoduln* nennen wollen.

Ein Formenmodul ist ein System algebraischer Formen (d. h. homogener ganzer Funktionen von n Variablen $x_1, x_2, \ldots, x_n$), welches die Eigenschaft hat, daß Summe, Differenz und Produkt zweier Formen des Systems immer wieder dem System angehören, aber auch das Produkt einer Form des Systems mit irgendeiner Form, die nicht zum System gehört, wieder eine Form des Systems bildet.

Der erste grundlegende Satz ist dann folgender: Aus jedem Formenmodul läßt sich eine endliche Anzahl m von Formen $F_1, F_2, \ldots, F_m$ auswählen, so daß jede Form von der Gestalt

$$F = A_1 F_1 + A_2 F_2 + \cdots + A_m F_m$$

(wo $A_1, A_2, \ldots, A_m$ beliebige Formen bedeuten, die ein homogenes F liefern), aber auch keine andere Form dem Modul angehört. Nach H. Weber (*Algebra* II, § 41) heißt der Inbegriff dieser m Formen $(F_1, F_2, \ldots, F_m)$ eine *Basis* des Formenmoduls.

Der Ausdruck Modul rührt bei Kronecker daher, daß er zwei Formen Φ, Φ' kongruent nach einem Modul nennt, in Zeichen

$$\Phi \equiv \Phi', \text{ Mod } (F_1, \ldots, F_m),$$

wenn Φ von der Form

$$\Phi = \Phi' + A_1 F_1 + \cdots + A_m F_m$$

ist.

Wird $F_1 = 0$, $F_2 = 0, \ldots, F_m = 0$, so wird auch $F = 0$. Aber es gehören nicht umgekehrt alle Formen, die mit der Basis $(F_1, F_2, \ldots, F_m)$ zugleich verschwinden, zu dem Formenmodul. Diese Formen bilden vielmehr einen neuen Modul, welcher der *Nullmodul* des ersten heißen kann. Es zeigt sich

jedoch, daß die Verschiedenheit eines Formenmoduls und seines Nullmoduls insofern eine beschränkte ist, als die Ordnung der Formen, die dem letzteren, aber nicht dem ersteren angehören, unter einer endlichen Grenze liegt. Die Basis des Nullmoduls besteht aus Formen niedrigster Zahl und Ordnung, deren Verschwinden das Verschwinden von $F_1, \ldots, F_m$ ersetzen kann.

Durch die Gleichungen $F_1 = 0, \ldots, F_m = 0$ werden gewisse μ Veränderliche $x_1, x_2, \ldots, x_\mu$ als algebraische Funktionen der $n - \mu$ übrigen, $x_{\mu+1}, \ldots, x_n$, festgelegt. Die Zahl μ heißt die *Stufe* des Formenmoduls. Ist die Stufenzahl $n - 1$, so gestatten die Gleichungen $F_i = 0$ nur eine endliche Zahl gemeinsamer Lösungen. Ist die Stufe n, so besteht der zugehörige Nullmodul aus sämtlichen homogenen Funktionen von $x_1, x_2, \ldots, x_n$ und für seine Basis können wir $(x_1, x_2, \ldots, x_n)$ wählen. Wir können ihn als den Fundamentalmodul bezeichnen.

Ist $F_1 = F_1'$ die Form höchster Ordnung o unter den Formen der Modulbasis, so lassen sich $r - 1$ Formen $F_2', F_3' \ldots F_r'$ gleicher Ordnung o bestimmen, derart, daß jede Form des Moduls in der Gestalt:

$$F = B_1 F_1' + B_2 F_2' + \cdots + B_r F_r'$$

darstellbar ist, wobei nunmehr auch $B_1, \ldots, B_r$ Formen von gleicher Ordnung bedeuten. Von den Formen

$$F' = \lambda_1 F_1' + \lambda_2 F_2' + \cdots + \lambda_r F_r'$$

(in welchem Ausdrucke $\lambda_1, \lambda_2, \ldots, \lambda_r$ irgendwelche Konstanten bedeuten) sagt man, sie bilden eine *Schar*. Jeder Formenmodul ist so mit einer Formenschar verknüpft, und diese Verknüpfung ist eine eindeutige, denn die Formenschar besteht aus allen Formen o^{ter} Ordnung, welche der Formenmodul enthält.

Ein einfaches Beispiel für einen Formenmodul bildet der durch die folgenden drei Formen gegebene:

$$F_1 = x_2 x_4 - x_3^2, \quad F_2 = x_2 x_3 - x_1 x_4, \quad F_3 = x_1 x_3 - x_2^2.$$

Diese drei Formen verschwinden, wenn

$$x_1 = \xi^3, \quad x_2 = \xi^2 \eta, \quad x_3 = \xi \eta^2, \quad x_4 = \eta^3$$

gemacht wird, was auch die Werte von ξ, η seien. Die sämtlichen Formen

$$F' = \lambda_1 F_1 + \lambda_2 F_2 + \lambda_3 F_3$$

bilden die zu dem Formenmodul gehörige Formenschar.

Im allgemeinen bestehen zwischen den Formen einer Modulbasis gewisse identische Beziehungen:

$$X_1^{(\varrho)} F_1 + X_2^{(\varrho)} F_2 + \cdots + X_m^{(\varrho)} F_m \equiv 0 \qquad (\varrho = 1, 2 \ldots q).$$

Die Koeffizienten A_i in der Darstellung einer beliebigen Form F des Moduls sind dann keineswegs vollkommen bestimmt, vielmehr lassen sie sich durch andere

$$A_i + \sum_{\varrho} \Gamma^{(\varrho)} X_i^{(\varrho)}$$

ersetzen, wo die $\Gamma^{(\varrho)}$ Formen bedeuten, die bis auf ihre Ordnung willkürlich bleiben, und die Summation über die Indizes ϱ zu erstrecken ist, für welche die Ordnung von $X_i^{(\varrho)} F_i$ die Ordnung von F nicht übersteigt.

In dem angeführten Beispiel sind die drei Formen F_1, F_2, F_3 durch die Beziehungen verknüpft:

$$x_1 F_1 + x_2 F_2 + x_3 F_3 = 0,$$
$$x_2 F_1 + x_3 F_2 + x_4 F_3 = 0,$$

und sonach wird identisch:

$$A_1 F_1 + A_2 F_2 + A_3 F_3 = A_1' F_1 + A_2' F_2 + A_3' F_3,$$

wenn

$$A_1' = A_1 + x_1 \Gamma' + x_2 \Gamma'',$$
$$A_2' = A_2 + x_2 \Gamma' + x_3 \Gamma'',$$
$$A_3' = A_3 + x_3 \Gamma' + x_4 \Gamma'',$$

und Γ', Γ'' willkürliche Formen bezeichnen, deren Ordnung um 1 niedriger ist als die der A.

Aus den Gleichungen

$$X_1^{(\varrho)} F_1 + \cdots + X_m^{(\varrho)} F_m = 0$$

folgt, daß

$$X_1 F_1 + \cdots + X_m F_m = 0$$

wird, wenn man

$$X_i = \sum_{\varrho} Y^{(\varrho)} X_i^{(\varrho)} \qquad (i = 1, 2, \cdots, m)$$

setzt. Bei dieser Darstellung der X_i sind aufs neue die $Y^{(\varrho)}$ im allgemeinen nicht vollständig bestimmt, vielmehr erhält man, wenn eine Lösung dieses Gleichungssystems für $Y^{(\varrho)}$ gefunden

ist, die allgemeinste Lösung, indem man die allgemeine Lösung des Gleichungssystems

$$\sum_{\varrho} Y^{(\varrho)} X_i^{(\varrho)} = 0 \qquad (i = 1, 2, \ldots, m)$$

hinzunimmt.

Alle Lösungen dieses Systems lassen sich durch gewisse s unter ihnen $Y_1^{(\varrho)}, Y_2^{(\varrho)}, \ldots, Y_s^{(\varrho)}$ linear ausdrücken, so daß man als Lösung

$$Y^{(\varrho)} = \sum_{\sigma} Z_\sigma^{(\varrho)} Y_\sigma^{(\varrho)} \qquad (\sigma = 1, 2, \ldots, s)$$

erhält. Aus diesen Gleichungen ergibt sich weiter ein Gleichungssystem:

$$\sum_{\sigma} Z_\sigma^{(\varrho)} Y_\sigma^{(\varrho)} = 0 \qquad (\varrho = 1, 2, \ldots, \eta)$$

usw. Die Kette dieser abgeleiteten Gleichungssysteme bricht aber, wie Hilbert gezeigt hat, nach höchstens m Schritten ab, indem spätestens das m^{te} Gleichungssystem keine Lösung mehr zuläßt.

Ist z. B. die Basis des Moduls (x_1, x_2, x_3, x_4), so lautet die erste Gleichung:

$$x_1 X_1 + x_2 X_2 + x_3 X_3 + x_4 X_4 = 0$$

und liefert sechs unabhängige Lösungen. Von diesen ist z. B. die erste $(X_1^{(1)}, X_2^{(1)}, X_3^{(1)}, X_4^{(1)}) = (x_2, -x_1, 0, 0)$. Damit erhält man für die $Y^{(\varrho)}$ das Gleichungssystem

$$\begin{aligned}
x_2 Y^{(1)} + x_3 Y^{(2)} + x_4 Y^{(3)} &= 0,\\
-x_1 Y^{(1)} + x_3 Y^{(4)} + x_4 Y^{(5)} &= 0,\\
-x_1 Y^{(2)} - x_2 Y^{(4)} + x_4 Y^{(6)} &= 0,\\
-x_1 Y^{(3)} - x_2 Y^{(5)} - x_3 Y^{(6)} &= 0.
\end{aligned}$$

Das zweite Gleichungssystem wird dann

$$\begin{aligned}
x_2 Z_1 - x_1 Z_2 = 0, \quad x_4 Z_3 - x_3 Z_4 = 0,\\
x_3 Z_1 - x_1 Z_3 = 0, \quad x_2 Z_4 - x_4 Z_2 = 0,\\
x_4 Z_1 - x_1 Z_4 = 0, \quad x_3 Z_2 - x_2 Z_3 = 0,
\end{aligned}$$

und seine allgemeine Lösung ist

$$Z_1 = x_1 U, \; Z_2 = x_2 U, \; Z_3 = x_3 U, \; Z_4 = x_4 U,$$

so daß das vierte Gleichungssystem

$$x_1 U = 0,\ x_2 U = 0,\ x_3 U = 0,\ x_4 U = 0$$

keine Lösung $V \neq 0$ zuläßt Die geometrische Deutung dieses Beispiels wird sehr einfach, wenn x_1, x_2, x_3, x_4 als homogene Koordinaten eines Punktes P im Raume interpretiert werden. X_1, X_2, X_3, X_4 sind dann die Koordinaten einer Ebene, die P enthält, $Y^{(1)}, Y^{(2)}, \ldots$ die Koordinaten einer geraden Linie, die durch P hindurchgeht und $Z_1, Z_2, \ldots$ die Koordinaten eines Punktes, der mit P zusammenfällt.

Ein Modul $\mathfrak{M}$ heißt durch einen Modul $\mathfrak{N}$ teilbar und dieser ein Teiler jenes, wenn jede Form, die zu $\mathfrak{M}$ gehört, auch in $\mathfrak{N}$ enthalten ist. Z. B. ist $(F_1, F_2, \ldots, F_m)$ die Basis eines Teilers des durch $(F_1, \ldots, F_r)$ bestimmten Moduls, wenn $r < m$ ist.

Der größte gemeinschaftliche Teiler zweier Moduln $\mathfrak{M}$, $\mathfrak{M}'$ ist der Modul, der durch Addition der sämtlichen Formen von $\mathfrak{M}$ zu sämtlichen Formen von $\mathfrak{M}'$ entsteht. Dieser besteht, wenn $(F_1, F_2, \ldots, F_m)$ die Basis von $\mathfrak{M}$, $(F_1', \ldots, F_{m'}')$ die Basis von $\mathfrak{M}'$ ist, aus allen Formen, die in der Gestalt

$$F = A_1 F_1 + \cdots + A_m F_m + A_1' F_1' + \cdots + A_{m'}' F_{m'}'$$

darstellbar sind.

Das kleinste enthaltende Vielfache zweier Moduln $\mathfrak{M}$ und $\mathfrak{M}'$ besteht aus allen Formen, die gleichzeitig in $\mathfrak{M}$ und $\mathfrak{M}'$ enthalten sind.

Die Zahl der Bedingungen, die eine Form F von genügend hoher Ordnung R erfüllen muß, damit sie zu einem Modul $\mathfrak{M}$ gehört, bezeichnet Hilbert, indem er sie als Funktion von R betrachtet, mit $\chi(R)$ und nennt sie die *charakteristische Funktion* des Moduls. $\chi(R)$ ist eine ganze rationale Funktion von R mit ganzzahligen Koeffizienten und von einer Ordnung d, die kleiner als die Zahl der in den Formen auftretenden Variablen ist. Diese Funktion wird, durch Binomialkoeffizienten $\binom{R}{i}$ ausgedrückt, in folgender Weise geschrieben:

$$\chi(R) = \chi_0 + \chi_1 \binom{R}{1} + \chi_2 \binom{R}{2} + \cdots + \chi_d \binom{R}{d}.$$

$\chi_0, \chi_1, \cdots, \chi_d$ sind hierbei gewisse dem Modul eigentümliche ganze Zahlen und d ist die *Dimension* des Moduls. Für das erste von uns gewählte Beispiel wird

$$\chi(R) = 1 + 3R,$$

und für den Modul (x_1, x_2, x_3, x_4) ist $\chi(R) = 0$. In der Tat gehört *jede* Form von x_1, x_2, x_3, x_4 zu diesem Modul.

Der für uns wichtigste Satz über Formenmoduln wird gewonnen, wenn man den allgemeinen Begriff eines *Formensystems* einführt. Hierunter wird eine Gesamtheit von Formen einer bestimmten Anzahl $x_1, x_2, \ldots, x_n$ von Variablen verstanden, von der nur vorausgesetzt zu werden braucht, daß man von jeder einzelnen Form entscheiden kann, ob sie zu dem System gehört oder nicht. Dann lautet der allgemeine *Hilbertsche Satz:*

Jedes Formensystem gehört zu einem Formenmodul, d. h. es lassen sich aus dem System m Formen $F_1, \ldots, F_m$ auswählen, so daß jede weitere Form in der Gestalt:

$$F = A_1 F_1 + \cdots + A_m F_m$$

darstellbar ist.

Der Beweis geschieht durch vollständige Induktion, d. h. durch den Schluß von $n-1$ auf n Variable. Um diesen Schluß möglich zu machen, wird zunächst die allgemeine Form F des Systems durch eine bestimmte Form F_1 desselben von der Ordnung r dividiert. Damit diese Division die Wirkung hat, daß sie die Ordnung von F für eine der Variablen unter die Ordnung r von F_1 herunterdrückt, wird zuerst die Substitution ausgeführt:

$$x_1 = x_1' + \lambda_1 y, \ldots, x_{n-1} = x_{n-1}' + \lambda_{n-1} y, \; x_n = y,$$

und die $\lambda_1, \ldots, \lambda_{n-1}$ werden so gewählt, daß das Glied mit y^r in F_1 nicht verschwindet. Dann wird die Division ausgeführt, indem man F und F_1 als Funktionen von y ansieht, und man erhält ein Resultat:

$$F = a_1 F_1 + \Phi_1,$$

wo in Φ_1 y höchstens zur $r-1^{\text{ten}}$ Potenz vorkommt. Für eine solche Form läßt sich durch den Schluß von $r-2$ auf $r-1$ das Theorem erweisen, indem man $\Phi_1 = y^{r-1}\varphi + \psi$ macht, wo φ y gar nicht mehr und ψ es höchstens zur $(r-2)^{\text{ten}}$ Ordnung enthält, und voraussetzt, das Theorem sei für $n-1$ Variable allgemein bewiesen (vgl. Webers *Algebra* II, § 41).

Gordan hat den Hilbertschen Satz bewiesen, indem er von den einzelnen Gliedern der Formen ausgeht (*Gött. Nachr.* (1899) 240, *Journ. de math.* (5) **6** (1900), 141).

Die Hauptanwendung, die Hilbert von seinem Satze gemacht hat, besteht in dem Nachweis der Endlichkeit des In-

variantensystems, d. h. des Satzes, daß sich alle Invarianten gegebener Formen durch gewisse unter ihnen ganz und rational ausdrücken lassen. Der große Vorzug dieses Beweises gegenüber den ursprünglichen Endlichkeitsbeweisen (Gordan, *Journ. f. Math.* **69** (1868), 343, *Math. Ann.* **2** (1870), 227, **5** (1872), 595) beruht außer der größeren Einfachheit darauf, daß er sich mühelos auf Formensysteme von beliebig vielen Veränderlichen übertragen läßt. Wir beschränken uns nur der Kürze wegen auf eine Binärform f mit den Koeffizienten $a_0, \ldots, a_n$.

Zunächst läßt sich nach dem Hilbertschen Theorem, da die Invarianten offenbar ein Formensystem der Koeffizienten $a_0, \ldots, a_n$ bilden, jede Invariante J mit Hülfe gewisser Invarianten $J_1, J_2, \ldots, J_\mu$ in der Form ausdrücken:

$$(\alpha) \qquad J = A_1 J_1 + A_2 J_2 + \cdots + A_\mu J_\mu.$$

Durch eine lineare Substitution $x_1 = \xi_1 y_1 + \eta_1 y_2$, $x_2 = \xi_2 y_1 + \eta_2 y_2$ gehe $a_0, \ldots, a_n$ über in $a_0', \ldots, a_n'$, und die den A_i entsprechenden Formen der $a_0', \ldots, a_n'$ seien A_i'. Dann besteht eine Gleichung von folgender Form:

$$\Delta^\gamma \cdot J = A_1' \Delta^{\gamma_1} \cdot J_1 + \cdots + A_\mu' \Delta^{\gamma_\mu} \cdot J_\mu,$$

wenn $\Delta = \xi_1 \eta_2 - \xi_2 \eta_1$ den Modul der Substitution bedeutet. Nach einem von Clebsch und Gordan herrührenden und von Mertens (*Wien. Ber.* **95** (1887), 942) zuerst allgemein angewandten Verfahren unterwerfe man nun die letzte Gleichung γ-mal dem Prozeß

$$\Omega = \frac{\partial^2}{\partial \xi_1 \partial \eta_2} - \frac{\partial^2}{\partial \xi_2 \partial \eta_1};$$

dann gehen die Potenzen von Δ in einfache Zahlfaktoren und die Koeffizienten A_i' in Invarianten über, so daß jetzt:

$$(\beta) \qquad J = J_1' J_1 + \cdots + J_\mu' J_\mu$$

wird und nun die Koeffizienten J_i' nicht bloß homogene Funktionen von $a_0, \ldots, a_n$, sondern Invarianten von f bedeuten. Wendet man auf diese dasselbe Verfahren an wie auf J und fährt so weiter fort, so gelangt man nach einer endlichen Anzahl von Schritten zu einer Darstellung von J als ganze rationale Funktion einer endlichen Zahl bestimmter Invarianten $J_1, \ldots J_\mu, \ldots J_\nu$, womit die Endlichkeit des Systems bewiesen ist.

Der Beweis läßt sich leicht auch auf das simultane System mehrerer Binärformen und die Kovarianten ausdehnen. Daß hier-

aus dann die Endlichkeit des Komitantensystems auch für Simultanformen folgt, hat Peano (*Torino Atti* **17** (1882) 73) gezeigt.

Der Hilbertsche Satz gestattet aber auch, den Beweis mit geringen Veränderungen auf Formen von beliebig vielen Variabeln auszudehnen. Er gibt ferner die Möglichkeit, die Endlichkeit des Systems der *Syzygien* oder identischen Relationen zwischen den Komitanten nachzuweisen.

L. Maurer (*Münch. Ber.* 1899, p. 147, *Math. Ann.* **57** (1903) 265) hat den Satz über die Endlichkeit des Invariantensystems in folgender allgemeinen Form bewiesen: alle ganzen rationalen Funktionen der Variabeln $x_1, \ldots, x_n$, die durch eine kontinuierliche Gruppe G in sich transformiert werden, lassen sich als ganze rationale Funktionen von einer begrenzten Anzahl unter ihnen darstellen. Dabei können noch $x_1, \ldots, x_n$ einem System algebraischer Gleichungen $\mathfrak{F}_1 = 0$, $\mathfrak{F}_2 = 0$, ... unterworfen werden, das gegenüber der Gruppe G invariant ist.

Die interessante Frage nach solchen Invarianten, von denen alle anderen rationale Funktionen mit ganzen Koeffizienten sind, ist nur in einem kleinen Aufsatze von H. Weber, *Gött. Nachr.* 1893, p. 109 berührt worden.

Von großer Wichtigkeit ist es, nicht bloß auf den rationalen Zusammenhang, sondern auch auf die algebraische Abhängigkeit der Invarianten einzugehen. Eine solche Abhängigkeit besteht immer zwischen $N - \nu + 2$ Formen, wenn N die Gesamtzahl der in den Grundformen vorhandenen Konstanten, ν die Konstantenzahl der Substitutionsgleichungen, nämlich bei m Veränderlichen m^2, bedeutet.

So ergibt sich bei einer Binärform zwischen $n-1$ Invarianten $(J, J_1, \ldots, J_{n-2})$ eine algebraische Gleichung $f(J, J_1, \ldots, J_{n-2}) = 0$. Hilbert hat nun nachgewiesen, daß diese Gleichung immer auf eine besondere Form gebracht werden kann. Zunächst nämlich kann man sich durch Erheben in geeignete Potenzen $J, \ldots, J_{n-2}$ auf denselben Grad gebracht denken. Dann kann man weiter lineare Funktionen von $J, J_1, \ldots, J_{n-2}$, etwa

$$J_1' = J_1 - \lambda_1 J, \ \ldots, \ J_{n-2}' = J_{n-2} - \lambda_{n-2} J,$$

an die Stelle von $J_1, \ldots, J_{n-2}$ selbst bringen und diese so bestimmen, daß in der Gleichung der Koeffizient des ersten Gliedes gleich 1 wird. Dies ist der Fall, wenn

$$f(1, \lambda_1, \ldots, \lambda_{n-2}) = 1$$

wird. So ergibt sich eine Gleichung von der Form:

$$(\gamma) \qquad J^k + \mathfrak{J}_1 J^{k-1} + \cdots + \mathfrak{J}_k = 0,$$

in der $\mathfrak{J}_1, \ldots, \mathfrak{J}_k$ ganze rationale Funktionen von $J_1', \ldots, J_{n-2}'$ sind. Hierbei läßt sich, wie Hilbert gezeigt hat, J so wählen, daß jede weitere Invariante J' eine rationale Funktion von $J, J_1', \ldots, J_{n-2}'$ wird.

Die Gesamtheit aller rationalen Funktionen von $J, J_1', \ldots, J_{n-2}'$ bildet aber einen *algebraischen Funktionenkörper*, und da hier allgemein die Homogenitätsbedingung hinzutritt, können wir von einem *algebraischen Formenkörper* reden. Alle Formen, ganze und gebrochene, dieses Körpers sind andererseits Invarianten, wenn man sich entschließt, zu den ganzen Invarianten gebrochene hinzuzunehmen, d. h. Brüche, deren Zähler und Nenner von gewöhnlichen (ganzen) Invarianten gebildet wird. Dieser Formenkörper heißt der Invariantenkörper; zu ihm gehören auch alle Konstanten, d. h. von den Koeffizienten der Grundformen unabhängigen Zahlen. Wir können aus ihm k Invarianten $\mathfrak{J}_1, \mathfrak{J}_2, \ldots, \mathfrak{J}_k$, wie z. B. $J^{k-1}, J^{k-2}, \ldots, J, 1$, auswählen, so daß jede Invariante $\mathfrak{J}$ sich in der Form darstellen läßt:

$$\mathfrak{J} = \mathfrak{A}_1 \mathfrak{J}_1 + \mathfrak{A}_2 \mathfrak{J}_2 + \cdots + \mathfrak{A}_k \mathfrak{J}_k,$$

wobei $\mathfrak{A}_1, \mathfrak{A}_2, \ldots, \mathfrak{A}_k$ rationale Funktionen von $J_1', \ldots, J_{n-2}'$ sind. Die Invarianten $\mathfrak{J}_1, \mathfrak{J}_2, \ldots, \mathfrak{J}_k$ bilden dann eine *Basis* des Invariantenkörpers.

Die Grundgleichung (γ) zeigt, daß wenn die Grundinvarianten $J_1', \ldots, J_{n-2}'$ verschwinden, auch J und damit jede ganze Invariante verschwindet. Es gilt aber auch der umgekehrte Satz, daß, wenn das Verschwinden von $n-2$ Invarianten $J_1', \ldots, J_{n-2}'$ das Verschwinden aller ganzen Invarianten zur Folge hat, jede solche Invariante J von genügend hohem Grade eine ganze algebraische Funktion von $J_1', \ldots, J_{n-2}'$ wird, d. h. einer Gleichung von der Form (γ) genügt. Von einem gewissen Grade an ist nämlich jede Form von der Gestalt (α) auch in der Gestalt

$$J = \mathfrak{A}_1 J_1' + \cdots + \mathfrak{A}_{n-2} J_{n-2}'$$

darstellbar, weil mit $J_1', \ldots, J_{n-2}'$ auch $J_1, \ldots, J_\mu$ verschwinden und deswegen die durch die Basen $(J_1', \ldots, J_{n-2}')$ und $(J_1, \ldots, J_\mu)$ festgelegten Moduln nur in den Formen, deren Grad unter einer gewissen Grenze S liegt, verschieden sein können. Die vorstehende Gestalt läßt sich aber wieder auf die folgende reduzieren:

$$J = \mathfrak{J}_1' J_1' + \cdots + \mathfrak{J}_{n-2}' J_{n-2}',$$

wo $\mathfrak{J}_1', \ldots, \mathfrak{J}_{n-2}'$ Invarianten bedeuten. Diese können, wenn ihr Grad über der Grenze S liegt, weiter mit Hilfe der Invarianten $J_1', \ldots, J_{n-2}'$ reduziert, d. h. auf eine Form wie die vorstehende gebracht werden, so daß man schließlich zu einer Darstellung von J gelangt, die wie folgt aussieht:

$$J = R_1 j_1 + R_2 j_2 + \cdots + R_w j_w ,$$

indem $R_1, R_2, \ldots$ ganze Funktionen von $J_1', \ldots, J_{n-2}'$ und $j_1, \ldots, j_w$ Invarianten sind, deren Grad $< S$ ist, und wir können, indem wir nötigenfalls einzelne der $R = 0$ setzen, annehmen, es umfassen $j_1, \ldots, j_w$ ein volles System linear unabhängiger Invarianten aller Grade $< S$. Wenden wir dann dieselbe Darstellungsform auch auf $Jj_1, Jj_2, \ldots, Jj_w$ an, so ergeben sich w Gleichungen von der Form:

$$\begin{aligned}
Jj_1 &= R_{11} j_1 + R_{12} j_2 + \cdots + R_{1w} j_w , \\
Jj_2 &= R_{21} j_1 + R_{22} j_2 + \cdots + R_{2w} j_w , \\
&\cdots\cdots\cdots\cdots\cdots \\
Jj_w &= R_{w1} j_1 + R_{w2} j_2 + \cdots + R_{ww} j_w ,
\end{aligned}$$

und durch Elimination der j aus diesen Gleichungen folgt:

$$\begin{vmatrix}
R_{11} - J & R_{12} & \ldots & R_{1w} \\
R_{21} & R_{22} - J & \ldots & R_{2w} \\
\cdots & \cdots & \cdots & \cdots \\
R_{w1} & R_{w2} & \ldots & R_{ww} - J
\end{vmatrix} = 0 ,$$

d. h. eine Gleichung von der Form (γ) für J.

§ 6. Invariante Bildungen der Formen mit mehr als zwei Veränderlichen.

Im binären Gebiet kommen als Komitanten nur zwei Arten, Invarianten und Kovarianten, in Betracht. Bei Formen mit mehr als zwei Veränderlichen hat man aber noch andere Bildungen ins Auge zu fassen. Werden nämlich in der Matrix

$$\begin{Vmatrix}
x_1' & x_2' & \ldots & x_r' \\
x_1'' & x_2'' & \ldots & x_r'' \\
\cdots & \cdots & \cdots & \cdots \\
x_1^{(m)} & x_2^{(m)} & \ldots & x_r^{(m)}
\end{Vmatrix}$$

die einzelnen Zeilen einer homogenen linearen Substitution unterworfen, so transformieren sich auch die Determinanten der Matrix linear. Wird die Zahl m der Zeilen in dieser Matrix insbesondere $= r - 1$, so besteht dieselbe aus r Determinanten, die wir bezeichnen mit

$$u_1, u_2, \ldots, u_r,$$

und die lineare Transformation dieser Determinanten ist von der Art, daß der Ausdruck

$$u_1 x_1 + u_2 x_2 + \cdots + u_r x_r$$

den Charakter einer Kovariante hat, indem er sich mit dem Substitutionsmodul multipliziert, wie man sofort sieht, wenn man die Matrix durch eine Zeile

$$x_1 \; x_2 \; \ldots \; x_r$$

zu einer Determinante ergänzt und bedenkt, daß diese Determinante in der Tat bei einer linearen Substitution sich mit dem Substitutionsmodul multipliziert. Die lineare Transformation, welche die Größen $u_1, \ldots, u_r$ erfahren, nennt man die *reziproke* zu der Transformation der Variablen $x_1, \ldots, x_r$ und bezeichnet die Variablensysteme $u_1, \ldots, u_r$ und $x_1, \ldots, x_r$ als *kontragredient*, während man Variablensysteme, die derselben Transformation unterliegen, als *kogredient* und solche, die unabhängigen linearen Transformationen unterworfen werden, als *digredient* charakterisiert. Zu beachten ist, daß die Determinanten einer Matrix mit $r - 1$ Reihen von je r kontragredienten Variablen wieder kogrediente Größen sind und daß die Differentiationssymbole

$$\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \cdots \frac{\partial}{\partial x_r}$$

sich formal wie kontragrediente Variable verhalten, was sofort zu erkennen ist, wenn man die Differentiation an der linearen Form $u_1 x_1 + \cdots + u_r x_r$ ausführt, wo sie unmittelbar die kontragredienten Größen $u_1, \ldots, u_r$ liefert.

Die Dimensionen, zu denen eine Form die kontragredienten Variablen enthält, wollen wir der größeren Deutlichkeit wegen als *Klasse*, nicht als Ordnung bezeichnen.

Wenn man nun die allgemeinsten invarianten Systeme sucht, so hat man eine beliebige Anzahl Grundformen mit beliebig vielen Reihen von r Veränderlichen vorauszusetzen. Diese Ver-

änderlichen kann man von vornherein als kogredient voraussetzen, indem man etwa vorkommende kontragrediente Variablenreihen sofort durch die Determinante einer $(r-1)$-reihigen Matrix ersetzt. Nun hat Capelli, ein Theorem von Clebsch (*Gött. Abh.* 1872, Math. Klasse, p. 3) modifizierend, gezeigt, daß sich der allgemeine Fall auf den reduzieren läßt, wo nur $r-1$ Reihen kogredienter Variablen vorhanden sind (*Acc. dei Lincei, Rendic.* (4) 7 (1891), 161).

Neben den Kovarianten hat man aber auch solche Bildungen zu betrachten, deren Variable den reziproken Transformationen unterworfen werden, also zu denen der Grundformen kontragredient sind. Diese Bildungen heißen *Kontravarianten* (Sylvester) oder zugehörige Formen (Aronhold). Bildungen, die beide Arten von Variablenreihen gleichzeitig enthalten, heißen *Divarianten* (Salmon) oder Zwischenformen (Aronhold).

Der Ausdruck $u_1x_1 + \cdots + u_rx_r$ heißt die *identische Kovariante*, und es gilt der Satz: Wenn eine Divariante nicht die Koeffizienten der Grundform, sondern nur eine Reihe kogredienter und eine Reihe kontragredienter Variablen enthält, ist sie notwendigerweise eine Potenz der identischen Kovariante.

Die wichtigsten invarianten Prozesse lassen sich aus den bei den Binärformen besprochenen sofort herleiten. Es ist zunächst der *Aronholdsche Prozeß*. Wir schreiben zwei Formen n^{ter} Ordnung f, g von r Variablen in der Gestalt:

$$f = \sum \binom{n}{\lambda\mu\ldots} a_{\lambda\mu\ldots} x_1^{\lambda} x_2^{\mu} \ldots, \qquad g = \sum \binom{n}{\lambda\mu\ldots} b_{\lambda\mu\ldots} x_1^{\lambda} x_2^{\mu} \ldots$$

und bezeichnen mit Q irgendeine Komitante, dann ist der Aronholdsche Prozeß durch die Gleichung definiert:

$$\delta Q = \sum_{\lambda\mu\ldots} b_{\lambda\mu\ldots} \frac{\partial Q}{\partial a_{\lambda\mu\ldots}}.$$

Die *Polarenbildung* oder der *Plueckersche Prozeß* wird durch die Gleichung

$$\Delta_y f = \sum_{i=1}^{r} \frac{\partial f}{\partial x_i} y_i$$

definiert, und der *Cayleyschen Prozeß*, ausgeführt an einer Form F, die r Variablenreihen $x_1' \ldots x_r'$, $x_1'' \ldots x_r''$, $\ldots$, $x_1^{(r)} \ldots x_r^{(r)}$ zu den Ordnungen $n', n'', \ldots, n^{(r)}$ enthält, wird:

$$\Omega F = \frac{1}{n' n'' \dots n^{(r)}} \begin{vmatrix} \frac{\partial F}{\partial x_1'} & \frac{\partial F}{\partial x_2'} & \dots & \frac{\partial F}{\partial x_r'} \\ \frac{\partial F}{\partial x_1''} & \frac{\partial F}{\partial x_2''} & \dots & \frac{\partial F}{\partial x_r''} \\ \dots & \dots & \dots & \dots \\ \frac{\partial F}{\partial x_1^{(r)}} & \frac{\partial F}{\partial x_2^{(r)}} & \dots & \frac{\partial F}{\partial x_r^{(r)}} \end{vmatrix}.$$

Über die invarianten Prozesse vgl. man die Arbeiten von Capelli, *Fondamenti di una teoria generale delle forme algebriche, Acc. dei Lincei Mem.* (3) **12** (1882) 529, *Math. Ann.* **29** (1881) 331, **37** (1890) 1, *Napoli Rendic.* (2) **7** (1893) 29, 155, *Giorn. di Mat.* **32** (1894) 376.

Unter den speziellen invarianten Bildungen für Formen mit beliebig vielen Veränderlichen ist zunächst die *Jacobische* oder *Funktionaldeterminante* zu nennen (vgl. S. 156). Sie setzt r Formen $f_1, f_2, \dots, f_r$ mit r Veränderlichen voraus, die für diese der Reihe nach von den Ordnungen $n_1, n_2, \dots, n_r$ sind, und wird durch den Ausdruck gegeben:

$$\Delta = \frac{1}{n_1 n_2 \dots n_r} \begin{vmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \dots & \frac{\partial f_1}{\partial x_r} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \dots & \frac{\partial f_2}{\partial x_r} \\ \dots & \dots & \dots & \dots \\ \frac{\partial f_r}{\partial x_1} & \frac{\partial f_r}{\partial x_2} & \dots & \frac{\partial f_r}{\partial x_r} \end{vmatrix}$$

(Jacobi, *Journ. f. Math.* **22** (1841) 319, *Werke* **3**, 393, *Ostwalds Klassiker* Nr. 78). Bildet man von je r aus $r+1$ Formen die Funktionaldeterminante, und von je r unter den letzteren wieder die Funktionaldeterminanten, so sind diese den Urformen proportional. Vgl. Clebsch, *Journ. f. Math.* **69** (1868) 355. **70** (1869) 175, Rosanes, ebenda **75** (1873) 166, Pasch, ebenda **80** (1875) 177. Die Bedeutung der Funktionaldeterminante besteht darin, daß sie verschwindet, sowie zwischen den r Formen, auf die sie sich bezieht, eine identische Relation

$$\Phi(f_1, f_2, \dots, f_r) \equiv 0$$

besteht.

Sucht man die Bedingung dafür, daß die r Gleichungen:

$$f_1 = 0, f_2 = 0, \dots, f_r = 0$$

zusammen bestehen können, so hat man aus diesen Gleichungen die r Variablen zu eliminieren und erhält dann eine Gleichung $R = 0$ zwischen den Koeffizienten der r Formen. Die linke Seite dieser Gleichung ist die *Resultante* der r Formen. Einen expliziten Ausdruck für die Resultante dreier ternären Formen hat Gordan gegeben, *Math. Ann.* **50** (1897) 113.

Die *Hessesche Determinante* einer Form f n^{ter} Ordnung von $x_1, \ldots, x_r$ (vgl. S. 160) ist durch den Ausdruck gegeben:

$$H = \text{Det}\left|\frac{1}{n(n-1)}\frac{\partial^2 f}{\partial x_i \partial x_j}\right| \qquad (i, j = 1, 2 \ldots).$$

Für eine quadratische Form

$$f = \sum_i \sum_j a_{ij} x_i x_j$$

reduziert sie sich auf deren Diskriminante, nämlich die aus den Koeffizienten a_{ij} gebildete Determinante. Die Hessesche Determinante ist die Jacobische Determinante der abgeleiteten Formen:

$$f_1 = \frac{1}{n}\frac{\partial f}{\partial x_1}, \ldots, f_r = \frac{1}{n}\frac{\partial f}{\partial x_r}.$$

Die Resultante D dieser Formen heißt die *Diskriminante* der Grundform f. Dieselbe ist eine Invariante vom Grade $r(n-1)^{r-1}$. Die Diskriminante einer Ternärform hat Gordan untersucht, *Münch. Ber.* **17** (1887), 477.

An neueren Werken, welche die allgemeine Theorie der Formen behandeln, nennen wir noch Elliot, *Algebra of Quantics*, Oxford 1895, Andoyer, *Théorie des Formes*, Paris 1898. Zu vergleichen sind auch die betreffenden Abschnitte aus H. Weber, *Algebra* und A. Capelli, *Algebra complementare*.

Von besonderem Interesse unter den Formen mit mehr als zwei Variablen sind die *Ternärformen*, da sie einerseits noch eine formal einfache Ausgestaltung erlauben und andererseits, indem man die Variablen x_1, x_2, x_3 als Punktkoordinaten in einer Ebene deutet, in der ebenen Geometrie eine anschauliche Deutung finden. Wir wollen uns fortan auf sie beschränken. Eine Zusammenfassung ihrer Theorie nach den symbolischen Methoden findet man in Clebschs *Vorlesungen über Geometrie*, hggb. von Lindemann (Leipzig 1876). Eine interessante Monographie über sie hat Study geliefert (*Methoden zur Theorie der ternären Formen*, Leipzig 1889).

Die Komitanten der ternären Formen lassen sich in einer ähnlichen Weise symbolisch darstellen wie die Komitanten der Binärformen, indem man sie formal auf solche Komitanten einer Reihe von Linearformen f_x, g_x, h_x ... zurückführt, welche für die Koeffizienten einer jeden dieser Linearformen vom n^{ten} Grade sind. Es ist z. B.:

$$f_x = f_1 x_1 + f_2 x_2 + f_3 x_3$$

anzunehmen, und außerdem wird die abkürzende Bezeichnung

$$(fgh) = \begin{vmatrix} f_1 & g_1 & h_1 \\ f_2 & g_2 & h_2 \\ f_3 & g_3 & h_3 \end{vmatrix}$$

gewählt. Die kontragredienten Variablen werden durch die Festsetzung

$$u_1 = y_2 z_3 - y_3 z_2, \quad u_2 = y_3 z_1 - y_1 z_3, \quad u_3 = y_1 z_2 - y_2 z_1$$

eingeführt, aus der

$$u_x = u_1 x_1 + u_2 x_2 + u_3 x_3 = (xyz)$$

folgt. Dies ist die *identische Kovariante*. Eine lineare Form der u wird mit u_a bezeichnet, sie läßt sich auch durch eine Determinante von der Gestalt (fgu) ersetzen. Die symbolischen Faktoren, die in einer Komitante vorkommen, lassen sich dann auf folgende vier Typen reduzieren:

1. Determinanten aus Koeffizienten der Linearformen vom Typus (fgh),
2. Linearformen der Variablen x vom Typus f_x,
3. Linearformen der Variablen u vom Typus (fgu),
4. die identische Kovariante u_x.

Für das Rechnen mit symbolischen Ausdrücken ist von Wichtigkeit der Satz, daß ein solcher Ausdruck, wenn er bei Vertauschung zweier gleichwertigen Symbole sein Vorzeichen ändert, identisch verschwindet. Ferner gelten die folgenden Identitäten, die, solange man

$$f_x, g_x, h_x, k_x, l_x, m_x$$

als wirkliche lineare Formen auffaßt, als einfache Determinantenrelationen anzusehen sind:

I. $(fgh)k_x - (ghk)f_x + (hkf)g_x - (kfg)h_x = 0.$

II. $(fgh)(klm) - (ghk)(flm) + (hkf)(glm) - (kfg)(hlm) = 0.$

III. $(fgu)h_x + (hfu)g_x + (ghu)f_x = (fgh)u_x.$

Von diesen drei Gleichungen sind die beiden letzten einfache Folgen der ersten. Hierzu kommt

IV. $$(fgu) = f_y g_z - g_y f_z,$$

wenn u_1, u_2, u_3 die oben angegebene Bedeutung haben und die den Multiplikationssatz der Determinanten ausdrückende Gleichung:

V. $$(fgh)(xyz) = \begin{vmatrix} f_x & g_x & h_x \\ f_y & g_y & h_y \\ f_z & g_z & h_z \end{vmatrix}.$$

Für die angeschriebenen Relationen ist es gleichgültig, welche Größen als konstant und welche als veränderlich angesehen werden. Man kann also aus ihnen neue Gleichungen ableiten, indem man die Koeffizienten der Linearformen $f, g, h \ldots$ durch kontragrediente Variablen $u, v, w \ldots$ und, wenn man will, x, y, z durch Konstante $\alpha, \beta, \gamma \ldots$ ersetzt. Man hat dabei eine Beziehung

$$x_1 = v_2 w_3 - v_3 w_2, \quad x_2 = v_3 w_1 - v_1 w_3, \quad x_3 = v_1 w_2 - v_2 w_1$$

vorauszusetzen. Auch kann man in der Gleichung II für $f, g, h, k \ldots$ Variable $x, y, z, t \ldots$ einsetzen, ebenso in der Gleichung I, indem man gleichzeitig für x umgekehrt f nimmt, so daß man die Gleichung erhält:

VI. $$(xyz) f_t - (yzt) f_x + (ztx) f_y - (txy) f_z = 0.$$

Damit aber sind die symbolischen Identitäten erschöpft, indem alle weiterhin auftretenden auf die gefundenen zurückgeführt werden können wie von Study, *Math. Ann.* **30** (1887) 120, bewiesen wurde. Eine Erweiterung dieses Satzes gab Pascal, *Acc. dei Lincei*, *Rendic.* (4) **4** (1888) 119, *Memorie* (4) **5** (1888) 375.

Clebsch hat gezeigt, daß die allgemeinsten ternären Systeme sich auf solche Grundformen zurückführen lassen, die nur zwei Reihen kontragredienter Variablen x_1, x_2, x_3 und u_1, u_2, u_3 enthalten. Eine derartige Form stellt nach Clebschs Ausdruck, gleich Null gesetzt, einen *Konnex* dar, und in diesen allgemeinsten Gebilden wollte Clebsch den Abschluß der algebraischen ebenen Geometrie erblicken (*Gött. Nachr.* 1872, S. 429, *Math. Ann.* **6**, 1873, S. 205. Man vgl. die Darstellung bei Clebsch-Lindemann, 1. Bd., 7. Abt.).

Eine Ternärform läßt sich im Gegensatz zu einer Binärform im allgemeinen nicht in Linearfaktoren zerlegen. Die

Bedingungen, unter denen dies eintritt, haben untersucht Brill, *Gött. Nachr.* 1893, 757, *Math. Ver.* **5** (1897) 52, *Math. Ann.* **50** (1898) 157, Gordan, ebenda **45** (1894) 410.

Von großer Bedeutung für die Theorie der ternären Formen ist das *Clebschsche Übertragungsprinzip,* welches dazu dient, aus bekannten binären Komitanten spezielle invariante Bildungen eines Systems ternärer Formen abzuleiten. Dies Prinzip hat, auf einem Gedanken von Hesse (*Journ. f. Math.* **66** (1866) 15, *Ges. Abhdlgn.* S. 531) fußend, Clebsch im *Journ. f. Math.* **59** (1861) 1 aufgestellt. Vgl. auch Gundelfinger, *Math. Ann.* **6** (1873) 16.

Man reduziert hierbei eine Ternärform, indem man in ihr

$$x_i = k_1 y_i + k_2 z_i \qquad (i = 1, 2, 3)$$

setzt und dann k_1, k_2 als Variable ansieht. So wird

$$f_x = k_1 f_y + k_2 f_z = k_1 \varphi_1 + k_2 \varphi_2 = \varphi_k,$$

und wenn man diese Gleichung symbolisch in die n^{te} Potenz erhebt,

$$f_x^n = \varphi_k^n.$$

Analog nehmen wir $g_x^n = \psi_k^n, \ldots$ an und lassen nachher $f_x^n, g_x^n, \ldots$ einerseits und $\varphi_k^n, \psi_k^n, \ldots$ anderseits zusammenfallen. Nun wird

$$(\varphi\psi) = \varphi_1\psi_2 - \varphi_2\psi_1 = f_y g_z - f_z g_y = (fgu) \text{ (nach IV).}$$

Eine Komitante der Binärform φ_k^n setzt sich aber aus symbolischen Faktoren vom Typus $(\varphi\psi)$ und φ_k zusammen. Wenn man dann für diese Faktoren die gefundenen Werte (fgu) und f_x einsetzt, ergibt sich eine Komitante der Ternärform $f = f_x^n$. So findet man die folgende formale Vorschrift, um aus dem symbolischen Ausdrucke einer Komitante einer Binärform eine Komitante (im allgemeinen eine Divariante) einer Ternärform gleicher Ordnung abzuleiten:

Man deute die symbolischen Faktoren f_x in jenem Ausdrucke auf drei Variable und hänge den Faktoren (fg) ein Symbol u an, so daß sie aus zweireihigen Determinanten zu dreireihigen Determinanten (fgu) werden. Der entstehende Ausdruck ist eine Divariante der Ternärform f, die aus der Binärform entsteht, indem man in ihrem symbolischen Ausdrucke f_x^n das f_x statt auf zwei auf drei Variable deutet.

Wir gehen nun zur Betrachtung der einfachsten Ternärformen über.

§ 7. Besondere ternäre Systeme.

Eine *quadratische Ternärform*

$$f = \sum a_{ij} x_i x_j \qquad (i, j = 1, 2, 3)$$

hat zwei invariante Bildungen: eine Invariante

$$D = 6 \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

und eine Kontravariante

$$F = -\tfrac{1}{2} \begin{vmatrix} a_{11} & a_{12} & a_{13} & u_1 \\ a_{21} & a_{22} & a_{23} & u_2 \\ a_{31} & a_{32} & a_{33} & u_3 \\ u_1 & u_2 & u_3 & 0 \end{vmatrix},$$

die wir, als quadratische Form der kontragredienten Variablen u, schreiben wollen:

$$F = \sum \alpha_{ij} u_i u_j \qquad (i, j = 1, 2, 3).$$

Ist nun eine zweite quadratische Form der x

$$f' = \sum a'_{ij} x_i x_j \qquad (i, j = 1, 2, 3)$$

gegeben, so erhalten wir auch für sie eine Invariante D' und eine Kontravariante:

$$F' = \sum \alpha'_{ij} u_i u_j \qquad (i, j = 1, 2, 3).$$

Der Aronholdsche Prozeß liefert dann zwei *simultane Invarianten* beider Formen:

$$E = \sum \frac{\partial D}{\partial a_{ij}} a'_{ij}, \qquad E' = \sum \frac{\partial D'}{\partial a'_{ij}} a_{ij}$$

und eine simultane (wiederum quadratische) Kontravariante:

$$H = \sum \frac{\partial F}{\partial a_{ij}} a'_{ij} = \sum \frac{\partial F'}{\partial a'_{ij}} a_{ij}.$$

Man erhält weiter zwei (bilineare) Divarianten

$$\mathfrak{B} = \frac{1}{4}\left(\frac{\partial F}{\partial u_1}\frac{\partial f'}{\partial x_1} + \frac{\partial F}{\partial u_2}\frac{\partial f'}{\partial x_2} + \frac{\partial F}{\partial u_3}\frac{\partial f'}{\partial x_3}\right),$$

$$\mathfrak{B}' = \frac{1}{4}\left(\frac{\partial F'}{\partial u_1}\frac{\partial f}{\partial x_1} + \frac{\partial F'}{\partial u_2}\frac{\partial f}{\partial x_2} + \frac{\partial F'}{\partial u_3}\frac{\partial f}{\partial x_3}\right),$$

und aus diesen zwei neue Divarianten

$$\mathfrak{C} = \frac{1}{2}\begin{vmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial \mathfrak{B}}{\partial x_1} & u_1 \\ \frac{\partial f}{\partial x_2} & \frac{\partial \mathfrak{B}}{\partial x_2} & u_2 \\ \frac{\partial f}{\partial x_3} & \frac{\partial \mathfrak{B}}{\partial x_3} & u_3 \end{vmatrix}, \qquad \mathfrak{C}' = \frac{1}{2}\begin{vmatrix} \frac{\partial f'}{\partial x_1} & \frac{\partial \mathfrak{B}'}{\partial x_1} & u_1 \\ \frac{\partial f'}{\partial x_2} & \frac{\partial \mathfrak{B}'}{\partial x_2} & u_2 \\ \frac{\partial f'}{\partial x_3} & \frac{\partial \mathfrak{B}'}{\partial x_3} & u_3 \end{vmatrix}.$$

Ersetzt man in $\mathfrak{B}$ x_i durch $\frac{1}{2}\frac{\partial F}{\partial u_i}$, so erhält man dasselbe, wie wenn man in $\mathfrak{B}'$ x_i durch $\frac{1}{2}\frac{\partial F'}{\partial u_i}$ ersetzt $(i = 1, 2, 3)$. Man bekommt so eine Kontravariante $\mathfrak{P}$ dritter Klasse, die das Eigentümliche hat, daß sie in drei Linearfaktoren zerfällt.

Wir finden weiter zwei Divarianten:

$$\varphi = \begin{vmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial f'}{\partial x_1} & u_1 \\ \frac{\partial f}{\partial x_2} & \frac{\partial f'}{\partial x_2} & u_2 \\ \frac{\partial f}{\partial x_3} & \frac{\partial f'}{\partial x_3} & u_3 \end{vmatrix}, \qquad \Phi = \begin{vmatrix} \frac{\partial F}{\partial u_1} & \frac{\partial F'}{\partial u_1} & x_1 \\ \frac{\partial F}{\partial u_2} & \frac{\partial F'}{\partial u_2} & x_2 \\ \frac{\partial F}{\partial u_3} & \frac{\partial F'}{\partial u_3} & x_3 \end{vmatrix}.$$

Wir können nun von F, F' ausgehend dieselben Bildungen wiederholen, die wir für f, f' aufgestellt haben. An die Stelle der Koeffizienten a_{ij}, a'_{ij} treten dann die α_{ij}, α'_{ij}, an die Stelle der x die u und umgekehrt. Die den D, D' entsprechenden Invarianten werden dabei bis auf einen Zahlfaktor deren Quadraten D^2 und D'^2 gleich. Die den F, F' korrespondierenden Kontravarianten $\mathfrak{f}, \mathfrak{f}'$ werden bis auf einen Zahlfaktor die Formen Df, $D'f'$, die $\mathfrak{B}$, $\mathfrak{B}'$ reproduzieren sich, φ, Φ gehen ineinander über, an die Stelle von H tritt:

$$\mathfrak{h} = \sum \frac{\partial \mathfrak{f}}{\partial \alpha_{ij}} \alpha'_{ij} = \sum \frac{\partial \mathfrak{f}'}{\partial \alpha'_{ij}} \alpha_{ij},$$

also eine quadratische simultane Kovariante der Urformen, und an die Stellen von $\mathfrak{C}$, $\mathfrak{C}'$ treten zwei Divarianten $\mathfrak{K}$, $\mathfrak{K}'$.

Anstatt der Kontravariante $\mathfrak{P}$ erhält man, indem man in $\mathfrak{B}'$ u_i durch $\frac{1}{2}\frac{\partial f}{\partial x_i}$ oder in $\mathfrak{B}$ u_i durch $\frac{1}{2}\frac{\partial f'}{\partial x_i}$ ersetzt, eine Kovariante Π, die ebenfalls in drei Linearfaktoren zerfällt. Nennt man diese Linearfaktoren a_x, b_x, c_x, und die von $\mathfrak{P}$ u_α, u_β, u_γ, so wird bis auf einen Zahlfaktor:

$$\mathfrak{P} = (bcu) \cdot (cau) \cdot (abu),$$
$$\Pi = (\beta\gamma x) \cdot (\gamma\alpha x) \cdot (\alpha\beta x),$$

gleichzeitig werden f, f' von der Form:

$$f = la_x^2 + mb_x^2 + nc_x^2, \quad f' = l'a_x^2 + m'b_x^2 + n'c_x^2,$$

und F, F' von der Form:

$$F = \lambda u_\alpha^2 + \mu u_\beta^2 + \nu u_\gamma^2, \quad F' = \lambda' u_\alpha^2 + \mu' u_\beta^2 + \nu' u_\gamma^2.$$

Das *System zweier quadratischer Ternärformen* haben behandelt Rosanes, *Math. Ann.* **6** (1873) 264, Gordan, *Math. Ann.* **19** (1882) 529, Perrin, *Bull. de la Soc. math. de France* **18** (1890) 1, Gerbaldi, *Annali di mat.* (2) **17** (1890) 161.

Das *System dreier quadratischer Ternärformen* haben untersucht Gundelfinger, *Journ. f. Math.* **80** (1875) 141, Ciamberlini, *Giorn. di mat.* **24** (1886) 141, Mertens, *Wien. Sitzungsber.* **93** (1886) 62, **99** (1890) 367, Gerbaldi, *Giorn. di Mat.* **27** (1889) 33, *Torino Atti* **25** (1890) 390, Fischer u. Mumelter, *Monatsh. f. Math.* **8** (1897) 97. Über die Resultante dreier quadratischen Ternärformen vgl. noch Sylvester, *Cambr. Dubl. Math. J.* **8** (1853) 256, *Werke* **1** 411, Cayley, *Journ. f. Math.* **57** (1860) 139, *Papers* **4**, 349, Gordan, *Journ. de math.* (5) **3** (1897) 97.

Die *kubische Ternärform* ist zuerst von Hesse bei seinen Untersuchungen über die Kurven dritter Ordnung (*Journ. f. Math.* **28**, 68, 97, **36**, 143, **38**, 241, **41**, 285, *Werke* 89, 123, 155, 193, 219) behandelt und das System ihrer invarianten Bildungen ist sodann von Aronhold in einer klassischen Arbeit (*Journ. f. Math.* **55** (1858) 93) so gut wie vollständig entwickelt worden. Es ist interessant, mit dieser Abhandlung die vollentwickelte Symbolik in den gewissermaßen abschließenden Arbeiten von Clebsch und Gordan über denselben Gegenstand, *Math. Ann.* **1** (1869) 57, **6** (1873) 436, zu vergleichen. Man sehe auch Gundelfinger, *Math. Ann.* **4** (1871) 144, **8** (1875) 136, Harnack, *Math. Ann.* **9** (1876) 218 und weiter Thomae, *Leipz. Ber.* **51** (1899) 317, Gordan, *Am. Math. Soc. Trans.* **1** (1900) 9, White, ebenda **1** (1900) 170, endlich in betreff des algebraischen Problems, das sich an die Wendepunkte der ebenen Kurve dritter Ordnung knüpft, Clebsch, *Journ. f. Math.* **58** (1861) 229, *Math. Ann.* **2** (1870) 382, Gundelfinger, *Math. Ann.* **5** (1872) 442.

Wir schreiben eine *kubische Ternärform*:

$$f = \sum_i \sum_j \sum_k a_{ijk} x_i x_j x_k \qquad (i, j, k = 1, 2, 3)$$

und setzen symbolisch:

$$f = f_x^3 = g_x^3 = h_x^3 = k_x^3 = f_x'^3 = g_x'^3 = h_x'^3.$$

Die Form hat dann zunächst zwei unabhängige Invarianten:

(a)
$$S = (fgh)(fgk)(fhk)(ghk),$$
$$T = (fgh)^2 (f'g'h)(f'h'g)(g'h'f)(f'g'h').$$

Aus diesen leitet man eine absolute Invariante $\frac{S^3}{T^2}$ ab und die Diskriminante

$$D = T^2 - \frac{1}{6} S^3.$$

Indem man die Linearform k_x durch u_x ersetzt, leitet man aus der Invariante S die Kontravariante ab:

(b)
$$\Sigma = (fgh)(fgu)(fhu)(ghu),$$

und, indem man h_x' durch u_x ersetzt, aus T:

(b′)
$$\mathsf{T} = (fgh)^2 (f'g'h)(fg'u)(f'gu)(f'g'u).$$

Nicht symbolisch ist:

(b″)
$$\Sigma = \tfrac{1}{4} \sum \frac{\partial S}{\partial a_{ijk}} u_i u_j u_k,$$
$$\mathsf{T} = \tfrac{1}{6} \sum \frac{\partial T}{\partial a_{ijk}} u_i u_j u_k.$$

Die Hessesche Kovariante

(c)
$$H = (fgh)^2 f_x g_x h_x$$

ist wie die Urform von der dritten Ordnung. Ist $S = 0$, so zerfällt H in drei Linearfaktoren. Wenn $T = 0$ ist, wird die Hessesche Kovariante von H bis auf einen Zahlfaktor wieder die Urform f. Nicht symbolisch ist, wenn

$$f_{ij} = \frac{1}{6} \frac{\partial^2 f}{\partial x_i \partial x_j}$$

gesetzt wird,

$$(c') \qquad H = 6 \begin{vmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{vmatrix}.$$

Wir nehmen

$$(c'') \qquad H = \sum_i \sum_j \sum_k b_{ijk} x_i x_j x_k .$$

Das Clebschsche Übertragungsprinzip liefert aus der quadratischen Kovariante einer kubischen Binärform die folgende Divariante der kubischen Ternärform:

$$(d) \qquad \Theta = (fgu)^2 f_x g_x ,$$

die man auf nichtsymbolischem Wege in folgender Weise gewinnt:

$$(d') \qquad \Theta = -2 \begin{vmatrix} f_{11} & f_{12} & f_{13} & u_1 \\ f_{21} & f_{22} & f_{23} & u_2 \\ f_{31} & f_{32} & f_{33} & u_3 \\ u_1 & u_2 & u_3 & 0 \end{vmatrix}$$

Setzt man nun weiter

$$\Theta_{ij} = \frac{1}{2} \frac{\partial^2 \Theta}{\partial x_i \partial x_j} ,$$

welche Größen die x nicht mehr enthalten, so bekommt man eine neue Kontravariante (6. Klasse)

$$(e') \qquad F = -2 \begin{vmatrix} \Theta_{11} & \Theta_{12} & \Theta_{13} & u_1 \\ \Theta_{21} & \Theta_{22} & \Theta_{23} & u_2 \\ \Theta_{31} & \Theta_{32} & \Theta_{33} & u_3 \\ u_1 & u_2 & u_3 & 0 \end{vmatrix},$$

die in symbolischer Darstellung lautet:

$$(e) \qquad F = (fgu)^2 (hku)^2 (fhu)(gku) .$$

Aus Θ läßt sich die Invariante S in folgender Weise ableiten:

$$(a') \qquad S = \frac{1}{16} \sum_{ij\mu\nu} \frac{\partial^4 \Theta}{\partial x_i \partial x_j \partial u_\mu \partial u_\nu} \cdot \frac{\partial^4 \Theta}{\partial x_\mu \partial x_\nu \partial u_i \partial u_j} .$$

Wir wenden nun den Aronholdschen Prozeß in der Weise an, daß wir ihn auf die Koeffizienten der Urform f und der

Hesseschen Kovariante H beziehen. Wir definieren ihn also dadurch, daß

$$\delta U = \sum \frac{\partial U}{\partial a_{ijk}} b_{ijk}$$

gesetzt wird. Dieser Definition nach ist

(f) $$\delta f = H.$$

Umgekehrt wird

(f′) $$\delta H = \tfrac{1}{2} S \cdot f$$

und weiter

(g) $$\delta \Sigma = 3\mathsf{T}, \qquad \delta \mathsf{T} = \tfrac{5}{6} S \cdot \Sigma.$$

(h) $$\delta S = 4T, \qquad \delta T = S^2.$$

Es geht also T aus S hervor wie folgt:

(h′) $$T = \frac{1}{4} \sum \frac{\partial S}{\partial a_{ijk}} b_{ijk}.$$

Die Diskriminante D läßt sich durch die Koeffizienten der Urform f und ihrer Hesseschen Form H unsymbolisch in sehr übersichtlicher Weise darstellen. Es wird:

$$D = 36 \begin{vmatrix} a_{111} & a_{122} & a_{133} & a_{123} & a_{131} & a_{112} \\ a_{211} & a_{222} & a_{233} & a_{223} & a_{231} & a_{212} \\ a_{311} & a_{322} & a_{333} & a_{323} & a_{331} & a_{312} \\ b_{111} & b_{122} & b_{133} & b_{123} & b_{131} & b_{112} \\ b_{211} & b_{222} & b_{233} & b_{223} & b_{231} & b_{212} \\ b_{311} & b_{322} & b_{333} & b_{323} & b_{331} & b_{312} \end{vmatrix}$$

Die Formeln (c′) (d′) (e′) (a′) (h′) (b″) bezeichnen den Weg, den man zu gehen hat, um die Kombinanten ohne Zuhilfenahme der Aronholdschen Symbolik zu gewinnen.

Wir vereinigen nun die Form f und ihre Hessesche H zu der Form

$$f_{\varkappa\lambda} = \varkappa f + \lambda H.$$

Alle Formen, die wir so erhalten, bilden das *syzygetische Formenbüschel* von f.

Wir suchen die Komitanten der Form $f_{\varkappa\lambda}$ auszudrücken durch die Komitanten von f und die Parameter $\varkappa, \lambda$. Zu dem Zwecke führen wir zunächst die folgenden Formen von $\varkappa, \lambda$ ein:

$$G_{\varkappa\lambda} = \varkappa^4 - S\varkappa^2\lambda^2 - \tfrac{4}{3}T\varkappa\lambda^3 - \tfrac{1}{12}S^2\lambda^4,$$
$$L_{\varkappa\lambda} = S\varkappa^2 + 2T\varkappa\lambda + \tfrac{1}{6}S^2\lambda^2,$$
$$M_{\varkappa\lambda} = \varkappa^3 + \tfrac{1}{2}S\varkappa\lambda^2 + \tfrac{2}{3}T\lambda^3,$$
$$N_{\varkappa\lambda} = 3\varkappa^2\lambda - \tfrac{1}{2}S\lambda^3,$$

zwischen denen die identische Beziehung besteht

$$G_{\varkappa\lambda} + \lambda^2 L_{\varkappa\lambda} = \varkappa M_{\varkappa\lambda} - \tfrac{1}{6}S\lambda N_{\varkappa\lambda}.$$

Dann werden die Invarianten von $f_{\varkappa\lambda}$

$$S_{\varkappa\lambda} = \frac{1}{2}\left(\frac{\partial L_{\varkappa\lambda}}{\partial\varkappa}M_{\varkappa\lambda} + \frac{\partial L_{\varkappa\lambda}}{\partial\lambda}N_{\varkappa\lambda}\right),$$
$$T_{\varkappa\lambda} = \frac{1}{16}\left[\frac{\partial G_{\varkappa\lambda}}{\partial\varkappa}\frac{\partial S_{\varkappa\lambda}}{\partial\lambda} - \frac{\partial G_{\varkappa\lambda}}{\partial\lambda}\frac{\partial S_{\varkappa\lambda}}{\partial\varkappa}\right],$$

und die Diskriminante

$$D_{\varkappa\lambda} = D \cdot G_{\varkappa\lambda}^3.$$

Die Hessesche Kovariante wird

$$H_{\varkappa\lambda} = \frac{1}{4}\left(\frac{\partial G_{\varkappa\lambda}}{\partial\varkappa}H - \frac{\partial G_{\varkappa\lambda}}{\partial\lambda}f\right),$$

und wir finden die Kontravarianten

$$\Sigma_{\varkappa\lambda} = M_{\varkappa\lambda}\Sigma + N_{\varkappa\lambda}\mathsf{T},$$
$$\mathsf{T}_{\varkappa\lambda} = \frac{1}{12}\left[\frac{\partial G_{\varkappa\lambda}}{\partial\varkappa}\cdot\frac{\partial\Sigma_{\varkappa\lambda}}{\partial\lambda} - \frac{\partial G_{\varkappa\lambda}}{\partial\lambda}\cdot\frac{\partial\Sigma_{\varkappa\lambda}}{\partial\varkappa}\right].$$

Besonders ausgezeichnet ist die biquadratische Form $G_{\varkappa\lambda}$. Wir finden ihre Invarianten:

$$i_g = 0, \qquad j_g = -\tfrac{1}{3}D.$$

Gleichzeitig wird die Diskriminante der quadratischen Form $L_{\varkappa\lambda}$ gleich $-2D$.

Die Hessesche Kovariante von $G_{\varkappa\lambda}$ ist

$$H_g = -\tfrac{1}{3}S_{\varkappa\lambda},$$

ihre Kovariante sechster Ordnung:

$$T_g = -\tfrac{1}{3}T_{\varkappa\lambda}.$$

Über die *biquadratische Ternärform* vgl. die Monographie von Pascal, *Napoli Atti* (2) **12** (1905) No. 13, woselbst auch die weitere Literatur angegeben ist.

Über *quaternäre Formen* vgl. man Mertens, *Wiener Sitzungsber.* **98** (1889) 691.

Zuletzt wollen wir die *kanonische Darstellung der Ternärformen* an zwei Beispielen erörtern.

Als die kanonische Darstellung einer kubischen Ternärform wird gewöhnlich die von Hesse (*Journ. f. Math.* **28** (1844) 68, *Werke* S. 115) herrührende Gestalt angesehen:

$$f = X_1^3 + X_2^3 + X_3^3 + 6kX_1X_2X_3.$$

Die Invarianten der Form sind dann

$$S = 24k(k^3 - 1), \qquad T = 6(8k^6 + 20k^3 - 1),$$

und die Diskriminante wird

$$D = 36(1 + 8k^3)^3.$$

Die Hessesche Kovariante lautet

$$H = -6[k^2(X_1^3 + X_2^3 + X_3^3) - (1 + 2k^3)X_1X_2X_3],$$

ferner wird in den kontragredienten Variabeln u_1, u_2, u_3

$$\Sigma = -6[k(U_1^3 + U_2^3 + U_3^3) + (1 - 4k^3)U_1U_2U_3],$$

$$\mathsf{T} = 2(10k^3 - 1)(U_1^3 + U_2^3 + U_3^3) + 4k^2(4k^3 + 5)U_1U_2U_3.$$

Um k zu bestimmen, setzt man

$$x = 6k^2,$$

dann ist x eine Wurzel der Gleichung vierten Grades:

$$x^4 - Sx^2 - \tfrac{4}{3}Tx - \tfrac{1}{12}S^2 = 0,$$

und zwar ergibt sich

$$x = \pm\sqrt{\tfrac{1}{6}(S + \sqrt[3]{6D})} \pm \sqrt{\tfrac{1}{6}(S + \varepsilon\sqrt[3]{6D})} \pm \sqrt{\tfrac{1}{6}(S + \varepsilon^2\sqrt[3]{6D})},$$

wenn ε eine primitive dritte Einheitswurzel bedeutet. Die Transformation auf die kanonische Form wird darauf gefunden aus der Identität

$$6(1 + 8k^3)X_1X_2X_3 = H + 6k^2f,$$

aus dieser folgt, daß X_1, X_2, X_3 den drei Linearfaktoren, in welche $H + 6k^2f$ zerfällt, proportional sind.

Für $S = 0$ läßt sich die kubische Form f auf die Summe dreier Kuben

$$f = X_1^3 + X_2^3 + X_3^3$$

reduzieren. Wenn $S \neq 0$, ist diese Darstellung unmöglich, wohl aber läßt sich die Form auf unendlich viele Arten als Summe von vier Kuben darstellen.

Eine Ternärform vierter Ordnung, die wir in ihrer allgemeinen Gestalt schreiben:

$$f = \sum_i \sum_j \sum_k \sum_l a_{ijkl} x_i x_j x_k x_l \qquad (i, j, k, l = 1, 2, 3)$$

läßt sich auf die Summe der vierten Potenzen von *fünf* Linearformen zurückführen, wenn die folgende Invariante verschwindet:

$$B = \begin{vmatrix} a_{1111} & a_{1122} & a_{1133} & a_{1123} & a_{1131} & a_{1112} \\ a_{2211} & a_{2222} & a_{2233} & a_{2223} & a_{2231} & a_{2212} \\ a_{3311} & a_{3322} & a_{3333} & a_{3323} & a_{3331} & a_{3312} \\ a_{2311} & a_{2322} & a_{2333} & a_{2323} & a_{2331} & a_{2312} \\ a_{3111} & a_{3122} & a_{3133} & a_{3123} & a_{3131} & a_{3112} \\ a_{1211} & a_{1222} & a_{1233} & a_{1223} & a_{1231} & a_{1212} \end{vmatrix}$$

Ist $B \neq 0$, so ist die Form nur auf die Summe von *sechs* vierten Potenzen zurückführbar. Man kann ihr aber auch die Gestalt geben:

$$f = \varphi(x_1^2, x_2^2, x_3^2) + x_1 x_2 x_3 x_4,$$

wo φ eine quadratische Form der dahinterstehenden Argumente und x_4 eine Linearform bedeutet. Nur in besonderen Fällen läßt sich f auf eine quadratische Form φ der Quadrate x_1^2, x_2^2, x_3^2 reduzieren.

Die Ternärform f vierter Ordnung besitzt eine ausgezeichnete Kovariante vierter Ordnung S, die Clebsch entdeckt und studiert hat, *Journ. f. Math.* **67** (1867) 360. Ist f auf eine lineare Verbindung

$$f = \sum_{i=1}^{5} a_i x_i^4$$

der vierten Potenzen von fünf Linearformen x_i zurückführbar, so wird S von der Form:

$$S = x_1 x_2 x_3 x_4 x_5 \sum_1^5 \frac{a_i}{x_i}.$$

Ist $f = \varphi(x_1^2, x_2^2, x_3^2)$, so wird S von der gleichen Form $S = \psi(x_1^2, x_2^2, x_3^2)$. S fällt nur dann mit f bis auf einen konstanten Faktor zusammen, wenn f ein vollständiges Quadrat oder von der Gestalt

$$f = x_1^3 x_2 + x_2^3 x_3 + x_3^3 x_1$$

ist. Diese besondere Form, die durch eine Gruppe von 168 linearen Transformationen in sich übergeht (vgl. S. 242), ist von Klein, *Math. Ann.* **14** (1879) 437 ff., *Theorie der elliptischen Modulfunktionen* I, S. 701 ff., und von Gordan, *Math. Ann.* **17** (1880) 217, 359, **20** (1882) 487, behandelt worden.

Kapitel VI.

Reihen, Produkte, Kettenbrüche.

Von *Paul Epstein* in Straßburg.

§ 1. Endliche Reihen.

Eine nach irgendeiner Vorschrift gebildete Folge von Zahlen

$$u_0,\ u_1,\ u_2,\ u_3, \cdots$$

heißt eine *Reihe*. Die Reihe heißt *endlich*, wenn ein letztes Glied u_n existiert, im andern Fall hat man eine *unendliche* Reihe.

Unter den endlichen Reihen sind die *arithmetischen* und *geometrischen* Reihen besonders hervorzuheben.

I. Eine *arithmetische Reihen k^{ter} Ordnung* liegt vor, wenn die *Glieder der k^{ten} Differenzenreihe konstant* und nicht null sind (vgl. Differenzenrechnung § 1).

Das *allgemeine Glied einer arithmetischen Reihe 1. Ordnung* ist also

$$u_\nu = a + \nu d. \qquad \nu = 0, 1, 2, \ldots n-1.$$

Die Summe einer arithmetischen Reihe 1. Ordnung von n Gliedern ist gleich dem Produkt der Gliederzahl mit dem arithmetischen Mittel des Anfangs- und Endglieds (Diophant).

Das allgemeine Glied u_ν einer arithmetischen Reihe k^{ter} Ordnung ist eine ganze rationale Funktion k^{ten} Grades von ν und jede Reihe mit dem allgemeinen Glied

$$u_\nu = c_0 + c_1\nu + c_2\nu^2 + \cdots + c_k\nu^k \qquad \nu = 0, 1, 2 \ldots$$

ist eine arithmetische Reihe k^{ter} Ordnung (de Lagny, Hist. de l'acad. d. sc. (1722) 281).

Anstatt der Potenzen von ν führt man zweckmäßig die *Binomialkoeffizienten* ein und hat

$$(1) \qquad u_\nu = a_0 + a_1\binom{\nu}{1} + a_2\binom{\nu}{2} + \cdots + a_k\binom{\nu}{k}.$$

Darin sind $a_0, a_1, a_2, , , .$ die *Anfangsglieder der* $0^{ten}, 1^{ten}, 2^{ten} \ldots$ *Differenzenreihe* $(a_0 = u_0)$.

Die Summe der n ersten Glieder der arithmetischen Reihe k^{ter} *Ordnung ist*

$$(2) \qquad n a_0 + \binom{n}{2} a_1 + \binom{n}{3} a_2 + \cdots + \binom{n}{k+1} a_k .$$

Die Formeln (1) und (2) sind von Jac. Bernoulli (1654 bis 1705) in der *Ars conjectandi* (1713) 98, (Ostwalds *Klassiker* Nr. 107) gegeben. Ihrem Inhalt nach waren sie schon Newton (*Principia* III^b, *Lemma* V) bekannt, vgl. auch Leibniz, *Math. Schriften* **1**, 27, herausg. von Gerhardt (Brief an Oldenburg vom 3. 2. 1673).

Zu den wichtigsten arithmetischen Reihen höherer Ordnung gehören die *Potenzsummen der n ersten natürlichen Zahlen.* Ist

$$S_k(n) = 1^k + 2^k + 3^k + \cdots + n^k, \quad S_0(n) = n,$$

so bestehen die Rekursionsformeln:

$$(n+1)^{k+1} = 1 + \binom{k+1}{1} S_k(n) + \binom{k+1}{2} S_{k-1}(n) + \cdots + \binom{k+1}{k} S_1(n) + S_0(n).$$

$$0 = n^{k+1} - \binom{k+1}{1} S_k(n) + \binom{k+1}{2} S_{k-1}(n) + \cdots + (-1)^{k+1} S_0(n).$$

Allgemein stellt sich $S_k(n)$ mit Hilfe der *Bernoullischen Zahlen*

$$B_1 = \frac{1}{2}; \; B_2 = \frac{1}{6}, \; B_4 = -\frac{1}{30}, \; B_6 = \frac{1}{42}, \; B_8 = -\frac{1}{30}, \cdots$$

$$B_3 = B_5 = B_7 = \cdots = 0$$

dar (vgl. Differenzenrechnung § 3), nämlich

$$S_k(n) = \frac{n^{k+1}}{k+1} + B_1 n^k + \frac{B_2}{2}\binom{k}{1} n^{k-1} + \frac{B_4}{4}\binom{k}{3} n^{k-3} + \frac{B_6}{6}\binom{k}{5} n^{k-5} + \cdots$$

oder symbolisch

$$S_k(n) = \frac{(n+B)^{k+1} - B^{k+1}}{k+1}.$$

Bis $S_{11}(n)$ sind die Potenzsummen bereits von Faulhaber (um 1612) berechnet worden. Die ersten von ihnen sind

$$S_1(n) = \frac{n(n+1)}{2}$$

$$S_2(n) = \frac{n(n+1)(2n+1)}{6}$$

$$S_3(n) = \frac{n^2(n+1)^2}{4}$$

$$S_4(n) = \frac{1}{30}n(n+1)(2n+1)(3n^2+3n-1)$$

$$S_5(n) = \frac{1}{12}n^2(n+1)^2(2n^2+2n-1)$$

$$S_6(n) = \frac{1}{42}n(n+1)(2n+1)(3n^4+6n^3-3n+1)$$

$$S_7(n) = \frac{1}{24}n^2(n+1)^2(3n^4+6n^3-n^2-4n+2).$$

Zwischen diesen Summen bestehen mannigfache Relationen, von denen außer der schon im 11. Jahrhundert den Arabern bekannten $S_3(n) = S_1(n)^2$ nur die Jacobische

$$S_7(n) + S_5(n) = 2S_1(n)^4$$

erwähnt sei. Lampe (*Journ. f. Math.* 84, 270 (1878)), vgl. Lucas, *Théorie des nombres*, p. 224, Bachmann, *Niedere Zahlentheorie* 2, 17 (1910).

Andere arithmetische Reihen sind:

Die Reihen der Potenzen der ungraden Zahlen:

$$1 + 3 + 5 + 7 + \cdots (2n-1) = n^2$$

$$1^2 + 3^2 + 5^2 + 7^2 + \cdots (2n-1)^2 = \frac{n(2n+1)(2n-1)}{3} = \binom{2n+1}{3}$$

$$1^3 + 3^3 + 5^3 + 7^3 + \cdots (2n-1)^3 = n^2(2n^2-1).$$

Allgemein:

$$1^k + 3^k + 5^k + \cdots + (2n-1)^k = S_k(2n) - 2^k S_k(n).$$

Die *Polygonalzahlen* bilden arithmetische Reihen 2. Ordnung, die *Pyramidalzahlen* solche 3. Ordnung (vgl. *Kombinatorik* S. 49). Die Summe der *Dreieckszahlen* ist

$$1 + 3 + 6 + 10 + 15 + \cdots + \frac{n(n+1)}{2} = \frac{n(n+1)(n+2)}{6},$$

die n^{te} *Tetraederzahl.*

II. Bei einer *geometrischen Reihe* haben je zwei aufeinanderfolgende Glieder ein *konstantes Verhältnis*. Eine solche Reihe von n Gliedern ist allgemein

$$a + aq + aq^2 + \cdots + aq^{n-1} = a\frac{q^n - 1}{q - 1}.$$

§ 2. Unendliche Reihen.

Bei einer unendlichen Reihe

$$u_1, u_2, u_3, \ldots$$

bilde man die *Summe der n ersten Glieder*

$$s_n = u_1 + u_2 + u_3 + \cdots + u_n$$

und untersuche den Grenzwert von s_n für $n = \infty$. Wenn ein solcher existiert und *endlich* ist, so heißt die Reihe *konvergent* und $\lim s_n = s$ heißt die *Summe* der Reihe; man schreibt

$$s = u_1 + u_2 + u_3 + \cdots.$$

Ist aber $\lim s_n = \pm\infty$, so ist die Reihe *divergent*, und wenn kein Grenzwert existiert, also obere und untere Grenze von s_n verschieden sind, so sagt man, die Reihe ist *unbestimmt* oder *oszilliert*.

Eine Reihe mit *komplexen* Gliedern ist konvergent, wenn die aus den reellen und die aus den imaginären Bestandteilen gebildete Reihe konvergiert.

Eine Reihe $\sum u_n$ ist immer konvergent, wenn die Reihe $\sum |u_n|$ der absoluten Werte ihrer Glieder konvergiert. In diesem Fall heißt die Reihe *absolut* konvergent; *sie konvergiert bei jeder beliebigen Anordnung ihrer Glieder und zwar stets gegen dieselbe Summe.* Deshalb nennt man eine derartige Reihe auch *unbedingt konvergent.*

Eine Reihe mit komplexen Gliedern ist *absolut konvergent*, wenn die beiden aus den reellen und aus den imaginären Teilen gebildeten Reihen absolut konvergieren.

Es kann aber auch eine Reihe $\sum u_n$ konvergieren, während die Reihe der absoluten Werte $\sum |u_n|$ *divergiert;* in diesem Falle heißt die ursprüngliche Reihe *einfach* konvergent, und ihre Summe hängt wesentlich von der Anordnung der Glieder ab. Bei einer solchen Reihe mit *reellen* Gliedern kann man durch geeignete Umstellung jeden beliebigen Wert erreichen. Deshalb nennt man eine derartige Reihe auch *bedingt konvergent.* Cauchy,

Résumé anal. (1833), 57; Dirichlet, *Berl. Abh.* (1837), 48, *Werke* **1**, 318; Riemann, *Werke*, S. 235; Schlömilch, *Z. f. Math.* **18**, 520 (1873), Pringsheim, *Math. Ann.* **23**, 455 (1883).

Eines der bekanntesten Beispiele einer bedingt konvergenten Reihe ist die Reihe $1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots$. Bei dieser Anordnung der Glieder ist ihre Summe $\ln 2$; läßt man aber auf je m positive Glieder n negative folgen, so ist die Summe $\ln 2 + \frac{1}{2} \ln\left(\frac{m}{n}\right)$. Es ist also z. B. $(m = 1,\ n = 4)$:

$$1 - \frac{1}{2} - \frac{1}{4} - \frac{1}{6} - \frac{1}{8} + \frac{1}{3} - \frac{1}{10} - \frac{1}{12} - \frac{1}{14} - \frac{1}{16} + \frac{1}{5} - \cdots = 0.$$

Das *notwendige und hinreichende Kriterium für die Konvergenz einer Reihe* ist von Bolzano (*Rein analytischer Beweis des Lehrsatzes usw.* 1817; Ostwalds *Klassiker* Nr. 153) gegeben worden, nämlich:

Die Reihe $u_1 + u_2 + u_3 + \cdots$ *ist konvergent, wenn für hinreichend große Werte von n der absolute Wert der Differenz*

$$|s_{n+p} - s_n| = |u_{n+1} + u_{n+2} + \cdots + u_{n+p}|$$

für jede natürliche Zahl p beliebig klein wird (vgl. S. 13).

Dieses Kriterium ist aber nur in wenigen Fällen wirklich anwendbar; man hat deshalb bequemer zu handhabende Kriterien aufgestellt, mit denen man in den meisten Fällen auskommt.

Wir erwähnen zunächst den einfachen Satz:

Wenn die Glieder einer Reihe abwechselnd positiv und negativ sind, ihrem absoluten Wert nach abnehmen und der Null zustreben, so konvergiert die Reihe (Leibniz *um* 1674, *Schriften* **5**, 112, *Brief an* Joh. Bernoulli *vom* 10. 1. 1714).

Im folgenden ist, wo nichts anderes bemerkt, von Reihen mit *reellen positiven* Gliedern die Rede.

Die Grundlage der meisten Kriterien bildet das Prinzip der *Reihenvergleichung*, indem man nämlich die vorgelegte Reihe Glied für Glied mit einer bereits als konvergent oder divergent erkannten Reihe mit reellen positiven Gliedern (*Majorante*) vergleicht.

Am häufigsten werden als Vergleichsreihen benutzt:

Die unendliche *geometrische* Reihe $1 + q + q^2 + \cdots$. Sie *konvergiert*, wenn $|q| < 1$, und ihre Summe ist dann $\frac{1}{1-q}$; sie *divergiert*, wenn $|q| \geqq 1$.

Die Reihen $1+\frac{1}{2^p}+\frac{1}{3^p}+\frac{1}{4^p}+\cdots$. Sie *konvergieren*, wenn $p>1$, *divergieren*, wenn $p\leqq 1$ ist (Waring, *Meditat. analyt.* 1781).

Die gebräuchlichsten Kriterien enthalten Ausdrücke, in denen entweder nur das allgemeine Glied u_n (*Kriterien erster Art*) oder das Verhältnis $\frac{u_{n+1}}{u_n}$ von zwei aufeinanderfolgenden Gliedern (*Kriterien zweiter Art*) vorkommt (Du Bois-Reymond, *J. f. Math.* **76**, 61 (1873)).

Es sei $\frac{1}{d_n}$ bzw. $\frac{1}{c_n}$ das allgemeine Glied einer divergenten bzw. konvergenten Reihe, ferner (a_n) eine ganz beliebige positive Zahlenfolge, so sind die *Hauptformen der Kriterien 1. und 2. Art* für die Divergenz und Konvergenz der Reihe $\sum u_n$ nach Pringsheim, *Math. Ann.* **35**, 297 (1890), *Enzykl.* **1**, 84:

Kriterien 1. Art:[1])	$\underline{\lim}\, d_n u_n > 0$	Divergenz.
	$\overline{\lim}\, c_n u_n < \infty$	Konvergenz.
Kriterien 2. Art:	$\overline{\lim}\left(d_n \frac{u_n}{u_{n+1}} - d_{n+1}\right) < 0$	Divergenz.
	$\underline{\lim}\left(a_n \frac{u_n}{u_{n+1}} - a_{n+1}\right) > 0$	Konvergenz.

Das letzte Kriterium wurde bereits von Kummer, *J. f. Math.* **13**, 171 (1835), für den speziellen Fall $a_n = n$ von Raabe, *J. f. Math* **11**, 309 (1834), angegeben und zur Transformation einer Reihe in eine stärker konvergierende benutzt (Kummer, *J. f. Math.* **16**, 206 (1837)); vgl. Catalan, *Mém. Belg. cour.* **33**, (1865). Durch geeignete Wahl der a_n, c_n, d_n kann man Skalen von immer wirksameren Kriterien bilden, unter denen die bekanntesten anderweitig aufgestellten Kriterien enthalten sind. Von diesen nennen wir:

Cauchys *Kriterium 1. Art (Anal. algebr.* (1821) 133):

$\overline{\lim}\sqrt[n]{u_n} > 1$ *Divergenz.* $\overline{\lim}\sqrt[n]{u_n} < 1$ *Konvergenz.*

Cauchys *Kriterium 2. Art* (ebda., jedoch bereits bei Waring, *Meditat. analyt.* (1781)):

---

1) $\overline{\lim}$ bedeutet den *oberen*, $\underline{\lim}$ den *unteren* Grenzwert.

$\lim \frac{u_{n+1}}{u_n} > 1$ *Divergenz*, $\lim \frac{u_{n+1}}{u_n} < 1$ *Konvergenz.*

Zu den Kriterien 1. Art gehört auch das *Integralkriterium von* Cauchy, *Anc. Exerc.* **2**, 221, (1827): *Ist* $f(x)$ *eine von einem bestimmten Wert* $x = a$ *ab positive, monoton bis zu null* (*für* $x = \infty$) *abnehmende Funktion, so konvergiert oder divergiert die Reihe* $\sum f(n)$, *je nachdem das Integral* $\int_a^\infty f(x)dx$ *einen Sinn hat oder nicht* (vgl. Hurwitz, *Math. Ann.* **44**, 83 (1894), Pringsheim, *Chicago Math. Congr. Papers* (1893) 328).

Die den oben angeführten Hauptformen entsprechenden Kriterien für Reihen mit *komplexen* Gliedern hat Pringsheim, *Arch. d. Math. u. Phys.* (3) **4**, 1 (1903) gegeben. Es sei darüber der Satz mitgeteilt:

Läßt sich bei einer Reihe $\sum u_n$ *eine Entwicklung*

$$\frac{u_n}{u_{n+1}} = 1 + \frac{a}{n} + \frac{r(n)}{n^2}$$

ansetzen, wobei $a = \alpha + i\beta$ *eine bestimmte, von n unabhängige Zahl,* $r(n)$ *eine mit wachsendem n endlich bleibende Größe bedeutet, so ist die Reihe unbedingt konvergent für* $\alpha > 1$, *divergent für* $\alpha \leqq 1$ (Gauß, *Werke* **3**, 139; Weierstraß, *Werke* **1**, 185). Die Reihe $\sum(-1)^n u_n$ ist für $\alpha > 1$ *unbedingt*, für $0 < \alpha \leqq 1$ *bedingt konvergent* und für $\alpha \leqq 0$ *divergent.*

Die folgenden Sätze gestatten, aus dem Verhalten einer oder mehrerer Reihen auf das Verhalten einer neu zu bildenden Reihe zu schließen:

1. *Ist* $s_n = u_1 + u_2 + \cdots + u_n$, *so konvergieren oder divergieren zu gleicher Zeit die Reihen* $\sum u_n$ *und* $\sum \frac{u_n}{s_n}$.

Ist $\sum u_n$ *divergent, so konvergiert für jedes* $\alpha > 0$ *die Reihe* $\sum \frac{u_n}{s_n^{1+\alpha}}$. *Vgl.* Abel, *J. f. Math.* **3**, 81 (1828), *Œuvres* **2**, 198; Dini, *Ann. Univ. Tosc.* **9**, 49 (1867).

2. *Sind* $a_1, a_2, a_3 \ldots$ *positive, mit wachsendem n nicht zunehmende Größen und* $\lim_{n=\infty} a_n = 0$ *und bleibt* $\sum_1^t b_n$ *für jedes t unter einer endlichen Schranke, so konvergiert die Reihe* $\sum a_n b_n$.

3. *Ist* $\sum a_n$ *absolut konvergent und bleibt* $\left|\sum_1^t b_n\right|$ *für jedes*

t unter einer endlichen Schranke, so ist die Reihe $\sum a_n b_n$ *absolut konvergent.*

4. *Ist* $\sum b_n$ *konvergent und* $\sum(a_{n+1} - a_n)$ *absolut konvergent, so konvergiert die Reihe* $\sum a_n b_n$ (Dedekind *in* Dirichlets *Vorl. üb. Zahlenth.*, 4. Aufl., 376, Hadamard, *Acta math.* **27**, 178 (1903)).

Die Beweise der letzten Sätze beruhen auf der von Abel, *J. f. Math.* **1**, 314 (1826) angegebenen *partiellen Summation*:

$$\sum_1^t a_n b_n = \sum_1^{t-1} (a_n - a_{n+1}) s_n + s_t a_t. \qquad (s_n = b_1 + b_2 + \cdots + b_n$$

Über das *Rechnen mit Reihen* gelten folgende Sätze:

1. Sind die Reihen $\sum u_n$ und $\sum v_n$ gegeben, so heißt $\sum(u_n + v_n)$ die *Summe* und $\sum(u_1 v_n + u_2 v_{n-1} + \cdots + u_n v_1)$ das *Produkt* der beiden Reihen.

2. *Die Summe von zwei konvergenten Reihen ist eine konvergente Reihe, deren Wert gleich der Summe der Werte der gegebenen Reihen ist.*

3. *Das Produkt von zwei konvergenten Reihen konvergiert, sobald eine der Reihen absolut konvergiert, und sein Wert ist gleich dem Produkt der Werte der beiden gegebenen Reihen* (Mertens, *J. f. Math.* **79**, 182 (1875), Jensen, *Nouv. Corr. math.* (1879) 430).

4. *Die Summe und das Produkt von zwei unbedingt konvergenten Reihen ist unbedingt konvergent* (Cauchy, *Cours d'analyse*).

Sind beide Reihen *bedingt* konvergent, so kann die Produktreihe divergieren, stets aber existiert ein Grenzwert für das Verhältnis $\frac{s_1 + s_2 + \cdots s_n}{n}$, und er ist gleich dem Produkt der Werte der beiden gegebenen Reihen (Cesàro, *Bull. d. sc. math.* **14**, 114 (1890), Borel, *Lec. s. l. sér. diverg.* 88. Weitere Literaturangaben bei Landau, *Rend. Circ. mat. Palermo* **24**, 81 (1907)).

Eine Reihe heißt *semikonvergent,* wenn die Summen

$$s_n = u_1 + u_2 + \cdots + u_n$$

bis zu einem bestimmten Wert N von n einer gewissen anderweitig definierten Zahl A immer näher kommen, für größere n sich aber immer weiter von A entfernen und schließlich unendlich groß werden. Diese Reihen sind also divergent, aber

sie verhalten sich bis zum N^{ten} Glied wie konvergente Reihen und sind in vielen Fällen zur numerischen Berechnung der Größe A sehr geeignet. Man wird auf sie geführt bei Anwendung der *Euler-Maclaurinschen Summenformel* (vgl. Differenzenrechnung), bei Ermittlung von *bestimmten Integralen* (vgl. Riemann-Weber, *Part. Differentialgl.*, Braunschweig 1900, S. 58), bei Auflösung von *linearen Differentialgleichungen* (vgl. Horn, *Gewöhnliche Differentialgl.*, Leipzig 1905, S. 188). In allen diesen Fällen ist A Funktion einer Veränderlichen x; dann sind die Glieder der Reihe ebenfalls Funktionen von x, und die Anwendbarkeit nimmt mit wachsendem x zu, d. h. die Zahl N der zu summierenden Reihenglieder ist von x abhängig, und die Differenz $A - s_n$ kann um so kleiner gemacht werden, je größer x ist. Deshalb dienen die semikonvergenten Reihen vor allem dazu, um die Werte einer Funktion für sehr große x zu berechnen. Die Reihe hat dann gewöhnlich die Form einer nach Potenzen von $\frac{1}{x}$ fortschreitenden Potenzreihe, und nach Poincaré (*Acta math.* 8, 285 (1886)) bezeichnet man eine divergente Entwicklung $a_0 + \frac{a_1}{x} + \frac{a_2}{x^2} + \cdots$ als *asymptotische Darstellung* einer Funktion $f(x)$, wenn für jedes n

$$\lim_{x=\infty} x^n \left[f(x) - \left(a_0 + \frac{a_1}{x} + \frac{a_2}{x^2} + \cdots + \frac{a_n}{x^n} \right) \right] = 0$$

ist. Vgl. Borel, *Leç. s. l. séries divergentes* 1901, p. 26.

Divergente Reihen wurden von fast allen Mathematikern des 18. Jahrhunderts unbedenklich benutzt; vor allem hat Euler von ihnen ausgedehnten Gebrauch in folgendem Sinn gemacht: Wenn sich eine Funktion $F(x)$ auf irgendeine Weise formal in eine Reihe von Funktionen

$$F(x) = \sum_{n=1}^{\infty} f_n(x)$$

entwickeln läßt, und die Reihe für $x = a$ divergiert, so wird der Funktionswert $F(a)$ als Summe der divergenten Reihe angesehen. So ist z. B.

$$1 - 1 + 1 - 1 + \cdots = \lim_{x=1} (1 - x + x^2 - \cdots) = \lim_{x=1} \frac{1}{1+x} = \frac{1}{2}.$$

Diese schon von Nik. Bernoulli, d'Alembert u. a. bekämpfte Auffassung (vgl. Pringsheim, *Enzykl.* 1, 105) wurde

seit Abel und Cauchy vollständig verlassen, und man arbeitete bis Ende des 19. Jahrhunderts nicht mehr mit divergenten Reihen. Erst in neuerer Zeit hat man mit Erfolg versucht, sie strenger Behandlung zugänglich zu machen. Den verschiedenen dabei vorgeschlagenen Methoden ist als leitender Gesichtspunkt gemeinsam, daß man einer divergenten Reihe einen bestimmten *konvergenten* Prozeß zuordnet. So ersetzen Laguerre (*Bull. Soc. math. France* **7**, 72 (1878/79) und Stieltjes (*Ann. fac. sc. Toulouse* **8**, (1894) und **9** (1895)) die divergenten Reihen durch Kettenbrüche und bestimmte Integrale, Padé (*Ann. éc. norm.* (3) **9**, (1892) supplém., *Acta math.* **18**, 97 (1894), *Ann. éc. norm.* (3) **19**, 187 (1902)) ebenfalls nach dem Vorgang von Laguerre durch konvergente Folgen von rationalen Funktionen, die als Näherungsbrüche von Kettenbrüchen aufgefaßt werden können. Von größerer Tragweite sind die *Methoden der Bildung von Mitteln*, die auf einen Satz von Frobenius, *J. f. Math.* **89**, 262 (1880) zurückgehen: *Ist* $a_0 + a_1 + \cdots + a_n = s_n$, *so ist*

$$\lim_{x=1} \sum_0^\infty a_n x^n = \lim_{n=\infty} \frac{s_0 + s_1 + \cdots s_n}{n+1},$$

sobald der Grenzwert auf der rechten Seite existiert und x wachsend den Grenzwert 1 *erreicht.*[1]) Es haben dann Cesàro, *Bull. sc. math.* **14**, 114 (1890) und Hölder, *Math. Ann.* **20**, 535 (1882) Mittel höherer Ordnung gebildet. Cesàro definiert

$$s_n^{(1)} = s_0 + s_1 + \cdots + s_n, \ldots s_n^{(k)} = s_0^{(k-1)} + \cdots + s_n^{(k-1)}$$

und bezeichnet die divergente Reihe $\sum a_n$ als γ-fach unbestimmt mit der Summe S, wenn der Grenzwert $\lim_{n=\infty} \frac{k!\, s_n^{(k)}}{n^k} = S$ für $k = \gamma$ existiert, für $k < \gamma$ aber nicht. Hölder bildet die Mittel

$$s_n^{(1)} = \frac{s_0 + \cdots + s_n}{n+1}, \ldots s_n^{(k)} = \frac{s_0^{(k-1)} + \cdots + s_n^{(k-1)}}{n+1}$$

und der divergenten Reihe $\sum a_n$ wird als Summe der Grenzwert $\lim_{n=\infty} s_n^{(k)} = S$ für das kleinste $k = \gamma$ zugeordnet. Für die Grenzwerte von Cesàro und Hölder gilt der Satz (Knopp, *Diss. Berlin* (1907), Schnee, *Math. Ann.* **67**, 110 (1909)), daß

1) Dieser Satz ist eine Verallgemeinerung des *Stetigkeitssatzes* von Abel (s. u. S. 434).

die Existenz des einen zugleich die des anderen mit demselben $k = \gamma$ nach sich zieht und dann beide denselben Wert S haben. Wichtige Anwendungen dieser Mittelwerte (bis zur 2. Ordnung) auf die Theorie der Fourierschen und Laplaceschen *Reihen* hat Fejér, *Math. Ann.* **58**, 51 (1904) und ebda. **67**, 76 (1909) gemacht.

Auch die Theorie von Borel (zusammenfassende Darstellung in *Leçons s. l. séries divergentes*, Paris 1901) beruht auf der Methode der Mittelwerte. Borel nennt eine Reihe $u_0 + u_1 + u_2 + \cdots$ *summierbar* und S ihre Summe, wenn der Grenzwert $S = \lim_{t=\infty} e^{-t} \sum_{n=0}^{\infty} s_n \frac{t^n}{n!}$ existiert. Er läßt sich dann durch das bestimmte Integral $S = \int_0^{\infty} e^{-t} u(t)\,dt$ ausdrücken, wobei $u(t) = \sum_0^{\infty} u_n \frac{t^n}{n!}$ die zur Reihe $\sum u_n$ *assoziierte ganze Funktion* ist. Notwendige Voraussetzung zur Summierbarkeit dieser Reihe ist also die Konvergenz von $u(t)$ für jedes endliche t. Borel definiert weiter den Begriff der *absolut* summierbaren Reihen und zeigt, daß sie in bezug auf Vertauschbarkeit der Glieder, Addition, Multiplikation genau die Eigenschaften der absolut konvergenten Reihen besitzen. Vgl. jedoch hierzu Hardy, *Quart. Journ.* **35**, 22 (1903).

§ 3. Reihen von Funktionen.

Sind die Glieder einer unendlichen Reihe Funktionen einer reellen Veränderlichen:

$$f_1(x) + f_2(x) + f_3(x) + \cdots \tag{1}$$

und ist

$$s_n(x) = f_1(x) + f_2(x) + \cdots f_n(x), \tag{2}$$

so heißt die Reihe *konvergent in einem Bereich* $a \leqq x \leqq b$, wenn für jedes x dieses Bereiches eine Zahl N gefunden werden kann, so daß zu jedem beliebigen positiven ε für jedes $n > N$ und jede natürliche Zahl p:

$$| s_{n+p}(x) - s_n(x) | < \varepsilon.$$

Der Wert von N hängt von der Größe ε ab und wird im allgemeinen in verschiedenen Punkten x des Bereiches verschie-

den sein. Wenn aber ein endlicher *von x unabhängiger Wert* der Zahl N gefunden werden kann, so heißt die Reihe *gleichmäßig konvergent* in dem Bereich $a \leqq x \leqq b$. Stokes, *Cambr. Trans.* 8, 533 (1847), Seidel, *Münch. Abh.* II. Kl. 7, 381 (1848). (Ostwalds *Klassiker* Nr. 116.)

Gleichmäßige und *absolute* Konvergenz sind nicht immer miteinander verbunden, man kann aber durch geeignete Zusammenfassung der Glieder eine gleichmäßig konvergierende Reihe auch absolut konvergent machen (Borel, *Acta math.* 24, 355). Es gilt ferner der Satz:

Eine Reihe von Funktionen konvergiert absolut und gleichmäßig, wenn die aus den oberen Grenzen der absoluten Beträge der Glieder gebildete Reihe konvergiert (Weierstraß, *Werke*, 2, 202).

An allen Stellen x, wo die Reihe (1) konvergiert, definiert sie eine Funktion $F(x) = \lim_{n=\infty} s_n(x)$. Zur Stetigkeit von $F(x)$ genügt es nicht, daß die Funktionen $f_1(x)$, $f_2(x)$, ... stetig sind. Es besteht jedoch der Satz:

Wenn eine Reihe von stetigen Funktionen in einem Bereich $a \leqq x \leqq b$ gleichmäßig konvergiert, so ist die Summe der Reihe in diesem Bereich eine stetige Funktion.

Es ist aber die gleichmäßige Konvergenz nicht *notwendige* Bedingung für die Stetigkeit der Reihensumme. Die notwendige und hinreichende Bedingung dafür wurde von Arzelà gegeben (*Mem. Acc. Bologna* (5) 8, 131 (1900); *Rend. Acc. Bologna* 7, (1903)). Vgl. Borel, *Leçons s. l. fonctions de var. réelles* (1905), 41.

Ihre wichtigste Anwendung finden die hier vorliegenden Reihen in der sogenannten *Darstellung willkürlicher Funktionen*, d. h. in dem Problem, eine gegebene Funktion durch eine Reihe darzustellen, die nach Funktionen von vorgeschriebener Natur fortschreiten. Dabei kommen hauptsächlich in Betracht: 1. *Entwicklungen nach ganzen rationalen Funktionen* (Polynomen), insbesondere *Potenzreihen*, 2. *Entwicklungen nach oszillierenden Funktionen*, insbesondere *Fouriersche (trigonometrische) Reihen*. Diese letzteren werden in Kap. XXII behandelt.

Die Entwicklungen nach Polynomen beruhen auf dem Satz von Weierstraß (*Berl. Sitzungsber.* (1885), 633, 789; *Werke* 3, 1):

Ist eine Funktion $f(x)$ in einem Intervall (a, b) mit Einschluß der Grenzen stetig und ε eine beliebige positive Zahl, so kann man eine ganze rationale Funktion $P(x)$ derart be-

stimmen, daß in dem ganzen Intervall

$$|f(x) - P(x)| < \varepsilon$$

ist.

Vgl. hierzu Borel, *Lec. s. l. fonctions de variables réelles*, Paris 1905, 50, woselbst zahlreiche Literaturangaben, Landau, *Rend. Circ. mat. Palermo* **25**, 337 (1908), Lebesgue, ebenda **26**, 325 (1908); Fejér, *Math. Ann.* **67**, 97 (1909).

Wählt man für ε der Reihe nach die Glieder $\varepsilon_1, \varepsilon_2, \varepsilon_3 \ldots$ einer konvergenten Reihe und ist $P_n(x)$ ein dem Wert ε_n entsprechendes Polynom, so ist

$$P_1 + (P_2 - P_1) + (P_3 - P_2) + \cdots + (P_n - P_{n-1}) + \cdots$$

eine Darstellung von $f(x)$ durch eine Reihe von Polynomen. Diese Reihe konvergiert absolut und gleichmäßig in dem ganzen Intervall (a, b).

In enger Beziehung mit dieser Darstellung steht das Problem der *Interpolation* durch ganze rationale Funktionen; dabei ist aber bemerkenswert, daß die Interpolationsformel von Lagrange (vgl. *Algebra* § 4, *Differenzenrechnung* § 2) bei Vermehrung der Argumente nicht unter allen Umständen eine immer stärkere Annäherung an die darzustellende Funktion liefert (vgl. Runge, *Ztschr. f. Math. u. Phys.* **46**, 229 (1901), Borel, *Leçons s. l. fonctions de var. réelles*, p. 74).

Tschebyscheff stellte sich die Aufgabe, eine beliebig gegebene Funktion $f(x)$ durch ein Polynom n^{ten} Grades $P_n(x)$ in einem Intervall so darzustellen, daß das Maximum von $|f(x) - P_n(x)|$ in dem Intervall möglichst klein wird. Über seine Arbeiten und die anschließende Literatur vgl. den Bericht von Burkhardt, *Entwicklungen nach oszillierenden Funktionen*, Leipzig 1908, S. 823—865; Liebmann, *Math. Ver.* **18**, 433 (1909).

Die wichtigsten polynomischen Entwicklungen sind die *Potenzreihen*, d. h. die Reihen von der Form

$$c_0 + c_1 x + c_2 x^2 + c_3 x^3 + \cdots.$$

Eine erschöpfende Behandlung erfordert die Heranziehung *komplexer* Werte für die Veränderliche x und ist Sache der Funktionentheorie. Hier sollen nur die folgenden Sätze erwähnt werden:

Eine Potenzreihe konvergiert absolut für alle Werte von x, deren absoluter Betrag kleiner ist als eine bestimmte positive

Zahl r, sie divergiert für alle x, deren absoluter Betrag größer ist als r.

Bei geometrischer Darstellung der komplexen Zahlen liegen alle Punkte x, für die die Reihe absolut konvergiert, innerhalb eines Kreises um den Nullpunkt mit dem Radius r. Dieser Kreis heißt *Konvergenzkreis*, r der *Konvergenzradius* der Potenzreihe.

Die Größe des Konvergenzradius wird durch den Satz von Cauchy (*Analyse algébr.* 1821, S. 286) bestimmt:

Der reziproke Wert von r ist gleich dem oberen Grenzwert der Zahlen $\sqrt[n]{|c_n|}$ $(n = 0, 1, 2, \ldots)$.

Ist dieser Grenzwert 0, so ist $r = \infty$, die Reihe konvergiert für *jeden* endlichen Wert von x; ist der Grenzwert ∞, so ist $r = 0$, die Reihe divergiert für jeden von Null verschiedenen Wert von x.

In jedem Bereich, welcher vollständig innerhalb des Konvergenzkreises liegt, also für $|x| \leq \varrho < r$ *konvergiert die Potenzreihe gleichmäßig und stellt daher dort eine stetige Funktion von x dar.*

Konvergiert die Reihe $\sum_0^\infty a_n x^n$ *für* $|x| < 1$ *und ist* $\sum_0^\infty a_n$ *konvergent, so ist* $\lim_{x=1} \sum a_n x^n = \sum a_n$. *Abelscher Stetigkeitssatz, J. f. Math.* **1**, 329 (1826).

Wenn zwei Potenzreihen für alle Werte von x in der Nachbarschaft von $x = 0$ *dieselbe Summe haben, so sind die Reihen identisch, d. h. es müssen die Koeffizienten gleicher Potenzen in beiden Reihen übereinstimmen.* Dieser Satz (Descartes 1637) spielt als *Methode der unbestimmten Koeffizienten* in älteren Arbeiten eine große Rolle. Er beruht auf der Stetigkeit einer konvergenten Potenzreihe.

Die Entwicklung einer gegebenen Funktion in eine Potenzreihe findet man, falls sie möglich ist, mit Hilfe des *Taylorschen Satzes*, vgl. *Differentialrechnung* § 6.

Als wichtigste Beispiele von Potenzreihen seien hervorgehoben:

1. *Binomische Reihe.*

$$(1+x)^n = 1 + nx + \frac{n(n-1)}{1 \cdot 2} x^2 + \frac{n(n-1)(n-2)}{1 \cdot 2 \cdot 3} x^3 + \cdots$$

$$= 1 + \binom{n}{1} x + \binom{n}{2} x^2 + \binom{n}{3} x^3 + \cdots.$$

Wenn n eine *ganze positive* Zahl ist, bricht die Reihe ab (Michael Stifel 1544); für jedes andere n ist die Reihe unendlich (Newton 1676) und *konvergiert absolut, sobald* $|x|<1$ *ist. Wenn* $|x|=1$ *ist*, konvergiert die Reihe nur für *positive* n absolut, für *negative* $n \leqq -1$ *divergiert* sie und für $-1<n\leqq 0$ ist sie *bedingt konvergent*, mit Ausnahme von $x=-1$.

Den ersten strengen Beweis des binomischen Satzes hat für *reelle* x und n erst Euler, *Petrop. Comment.* **19**, 103 (1775), für *komplexe* Werte von x Cauchy, *Analyse algébr.* 1821, gegeben, die abschließende, auch für die allgemeine Theorie der Reihen grundlegende Untersuchung aber, in der auch komplexe Werte von n herangezogen werden und der Grenzfall $|x|=1$ vollständig erledigt wird, bildet den Gegenstand der klassischen Abhandlung von Abel, *J. f. Math.* **1**, 311 (1826) (*Ostwalds Klassiker* Nr. 71).

Wichtige Spezialfälle der binomischen Reihe:

$$\frac{1}{1+x} = 1 - x + x^2 - x^3 + \cdots$$

$$\frac{1}{(1+x)^2} = 1 - 2x + 3x^2 - 4x^3 + \cdots$$

$$\sqrt{1+x} = 1 + \frac{1}{2}x - \frac{1\cdot 1}{2\cdot 4}x^2 + \frac{1\cdot 1\cdot 3}{2\cdot 4\cdot 6}x^3 - \frac{1\cdot 1\cdot 3\cdot 5}{2\cdot 4\cdot 6\cdot 8}x^4 + \cdots$$

$$\frac{1}{\sqrt{1+x}} = 1 - \frac{1}{2}x + \frac{1\cdot 3}{2\cdot 4}x^2 - \frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}x^3 + \frac{1\cdot 3\cdot 5\cdot 7}{2\cdot 4\cdot 6\cdot 8}x^4 + \cdots.$$

2. Die *Exponentialreihen:*

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$$

$$a^x = 1 + \frac{x \ln a}{1!} + \frac{x^2 (\ln a)^2}{2!} + \cdots.$$

Darin bedeutet

$$e = \lim_{n=\infty}\left(1+\frac{1}{n}\right)^n = 1 + \frac{1}{1!} + \frac{1}{2!} + \cdots$$

$$= 2{,}71828\ 18284\ 59045\ 23536\ 02874\ 71353\ldots$$

die *Basis der natürlichen Logarithmen.* Die Reihen konvergieren für jedes endliche x. Die Reihe für e^x wurde von Newton (*Opuscula* **1**, 20) durch Umkehrung der logarithmischen Reihe gefunden.

3. Die *logarithmischen Reihen.*

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots$$

konvergiert für $|x| \leqq 1$ mit Ausnahme von $x = -1$. Diese Reihe wurde als erstes Beispiel einer Potenzreihe (gleichzeitig mit der geometrischen Reihe für $\frac{1}{1+x}$) von Mercator 1668 gefunden. Bei $x = 1$ hat man die bedingt konvergente Reihe für $\ln 2$.

$$\frac{1}{2} \ln \frac{1+x}{1-x} = x + \frac{x^3}{3} + \frac{x^5}{5} + \cdots$$

konvergiert für $|x| \leqq 1$ mit Ausnahme von $x = \pm 1$. Daraus für $x = \frac{1}{2z+1}$

$$\ln(z+1) = \ln z + 2\left[\frac{1}{2z+1} + \frac{1}{3(2z+1)^3} + \frac{1}{5(2z+1)^5} + \cdots\right]$$

4. Die *goniometrischen Reihen.*

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots$$

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots$$

konvergieren für jedes endliche x. Sie sind bereits von Newton (*Opuscula* **1**, 24) aufgestellt worden. Euler (*Introd. in anal. infinit.* (1748) 104) benutzt zu ihrer Ableitung den Satz von Moivre (*Misc. analyt.* 1730):

$$(\cos x + i \sin x)^n = \cos nx + i \sin nx$$

und findet zugleich den Zusammenhang der goniometrischen Funktionen mit der Exponentialfunktion:

$$e^{ix} = \cos x + i \sin x.$$

5. Die folgenden Reihen sind mit den *Bernoullischen* und *Eulerschen Zahlen* gebildet. Näheres über diese siehe Kapitel XXII.

Die *Bernoullischen Zahlen*:

$$B_0 = 1, \quad B_1 = \frac{1}{2}, \quad B_3 = B_5 = B_7 = \cdots = 0$$

$$B_2 = \frac{1}{6}, B_4 = -\frac{1}{30}, B_6 = \frac{1}{42}, B_8 = -\frac{1}{30}, B_{10} = \frac{5}{66}, B_{12} = -\frac{691}{2730},$$

$$B_{14} = \frac{7}{6}, B_{16} = -\frac{3617}{510}, B_{18} = \frac{43867}{798}, B_{20} = -\frac{174611}{330}, \cdots$$

sind durch die symbolische Gleichung

$$(B+1)^n - B^n = n \qquad n = 1, 2, 3, \ldots$$

definiert. Ihre erzeugende Funktion ist

$$\frac{x}{e^x - 1} = \sum_{k=0}^{\infty} (-1)^k \frac{B_k}{k!} x^k$$

$$= 1 - \frac{x}{2} + \frac{B_2}{2!} x^2 + \frac{B_4}{4!} x^4 + \cdots,$$

konvergent für $|x| < 2\pi$.

$$x \operatorname{ctg} x = \sum_{k=0}^{\infty} (-1)^k \frac{2^{2k} B_{2k}}{(2k)!} x^{2k} \qquad \text{konvergiert für } |x| < \pi$$

$$\operatorname{tg} x = \sum_{k=0}^{\infty} (-1)^{k-1} \frac{2^{2k}(2^{2k}-1)}{(2k)!} B_{2k} x^{2k-1} \qquad \text{„} \qquad |x| < \frac{\pi}{2}$$

$$\frac{x}{\sin x} = 2 \sum_{k=0}^{\infty} (-1)^{k-1} \frac{2^{2k-1}-1}{(2k)!} B_{2k} x^{2k} \qquad \text{„} \qquad |x| < \pi$$

$$\ln \frac{\sin x}{x} = -\sum_{k=1}^{\infty} (-1)^{k-1} \frac{2^{2k-1} B_{2k}}{k \cdot (2k)!} x^{2k} \qquad \text{„} \qquad |x| < \pi$$

$$\ln \cos x = -\sum_{k=1}^{\infty} (-1)^{k-1} \frac{2^{2k-1}(2^{2k}-1)}{k \cdot (2k)!} B_{2k} x^{2k} \qquad \text{„} \qquad |x| < \frac{\pi}{2}$$

$$\ln \frac{\operatorname{tg} x}{x} = 2 \sum_{k=1}^{\infty} (-1)^{k-1} \frac{2^{2k-1}(2^{2k-1}-1)}{k \cdot (2k)!} B_{2k} x^{2k} \qquad \text{„} \qquad |x| < \frac{\pi}{2}$$

Die *Eulerschen Zahlen:*

$$E_0 = 1, \quad E_2 = -1, \quad E_4 = 5, \quad E_6 = -61, \quad E_8 = 1385,$$
$$E_{10} = -50521, \quad E_{12} = 2702765, \ldots$$
$$E_1 = E_3 = E_5 = \cdots = 0$$

sind durch die symbolische Gleichung

$$(E+1)^n + (E-1)^n = 0 \qquad n = 1, 2, 3, \ldots$$

definiert. Ihre erzeugende Funktion ist

$$\frac{2}{e^x + e^{-x}} = \sum_{k=0}^{\infty} \frac{E_k}{k!}\, x^k = \sum_{k=0}^{\infty} \frac{E_{2k}}{(2k)!}\, x^{2k},$$

oder, wenn xi an Stelle von x gesetzt wird:

$$\frac{1}{\cos x} = \sum_{k=0}^{\infty} (-1)^k \frac{E_{2k}}{(2k)!} x^{2k} \qquad |x| < \frac{\pi}{2}.$$

6. Die *zyklometrischen Reihen.*

$$\arcsin x = x + \frac{1}{2}\frac{x^3}{3} + \frac{1 \cdot 3}{2 \cdot 4}\frac{x^5}{5} + \cdots$$

konvergiert für $|x| \leqq 1$ und gibt für reelle x den zwischen $-\frac{\pi}{2}$ und $+\frac{\pi}{2}$ liegenden *Hauptwert* der arcsin-Funktion.

$$\operatorname{arctg} x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \cdots$$

konvergiert für $|x| \leqq 1$, ausgenommen $x = \pm i$, und gibt für reelle x den zwischen $-\frac{\pi}{4}$ und $+\frac{\pi}{4}$ liegenden *Hauptwert* der arctg-Funktion. Die Reihe wurde 1671 von Gregory gefunden. Für $x = 1$ gibt sie die bedingt konvergente, zur numerischen Berechnung von π nicht geeignete Entwicklung

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots.$$

Dagegen liefern die aus dem Additionstheorem

$$\operatorname{arctg} u + \operatorname{arctg} v = \operatorname{arctg} \frac{u + v}{1 - uv}$$

folgenden Relationen

$$\frac{\pi}{4} = 2 \operatorname{arctg} \frac{1}{3} + \operatorname{arctg} \frac{1}{7}$$

$$= 4 \operatorname{arctg} \frac{1}{5} - \operatorname{arctg} \frac{1}{239}$$

sehr schnell konvergierende Entwicklungen für π. Die erste ist von Vega, die zweite von Machin 1706 angegeben worden; mit ihr hat Shanks, *Proc. Roy. Soc.* **22**, 45 (1873) (vgl. *Ztschr. f. math. Unt.* **26**, 263 (1895)) die Zahl π bis auf 707 Stellen berechnet. Es ist

$\pi = 3{,}14159\ 26535\ 89793\ 23846\ 26433\ 83279\ 50\ldots.$

§ 4. Mehrfache Reihen.

Eine *Doppelreihe* entsteht durch Summation einer nach irgendeiner Vorschrift zu bildenden Folge von doppelt unendlich vielen Zahlen $u_{\mu\nu}\binom{\mu=1,2,3,\ldots}{\nu=1,2,3,\ldots}$. Ist

$$s_{m,n}=\sum_{\mu=1}^{m}\sum_{\nu=1}^{n}u_{\mu\nu},$$

so heißt die Doppelreihe *konvergent* und die endliche Zahl S ihre *Summe*, wenn stets

$$\lim_{\substack{m=\infty\\ n=\infty}} s_{m,n}=S$$

ist, wie auch immer m und n unabhängig voneinander ins Unendliche wachsen (vgl. S. 34). Man schreibt dann[1])

$$\sum_{\mu=1}^{\infty}{}'\sum_{\nu=1}^{\infty}u_{\mu\nu}=S.$$

Existiert aber kein bestimmter Grenzwert der $s_{m,n}$ oder ist $\lim s_{m,n}=\pm\infty$, so heißt die Doppelreihe *divergent*.

Aus den Gliedern der Doppelreihe kann man durch Heraushebung von einfach unendlichen Folgen beliebig viele einfache Reihen (Partialreihen) bilden. Zu diesen gehören die *Zeilenreihen* $\sum_{\nu=1}^{\infty}u_{\mu\nu}$ (mit festem μ) und die *Spaltenreihen* $\sum_{\mu=1}^{\infty}u_{\mu\nu}$ (mit festem ν). Mit der Konvergenz der Doppelreihe ist nicht durchaus diejenige aller Zeilen- oder Spaltenreihen verbunden (Stolz, *Math. Ann.* **24**, 157 (1884)), nur für Doppelreihen mit *positiven* Gliedern muß das der Fall sein, und wenn eine Doppelreihe sowie ihre sämtlichen Zeilen- und Spaltenreihen konvergieren, so ist

$$\sum_{\mu}\sum_{\nu}u_{\mu\nu}=\sum_{\mu}\Big(\sum_{\nu}u_{\mu\nu}\Big)=\sum_{\nu}\Big(\sum_{\mu}u_{\mu\nu}\Big).$$

---

1) Diese Schreibweise soll nicht bedeuten, daß zuerst nach ν und dann nach μ zu summieren ist. Soll dies der Fall sein, so schreibt man $\sum_{\mu=1}^{\infty}\Big(\sum_{\nu=1}^{\infty}u_{\mu\nu}\Big)$. Dies ist in Wirklichkeit keine Doppelreihe, sondern eine einfache Reihe, gebildet mit den Summen der (als konvergent vorauszusetzenden) Reihen $\sum_{\nu}u_{\mu\nu}$.

Man kann auf beliebig viele Arten die Glieder einer Doppelreihe so anordnen oder in Gruppen zusammenfassen, daß eine einfache Reihe entsteht. Diese einfachen Reihen brauchen bei Konvergenz der Doppelreihe nicht sämtlich zu konvergieren. Besonders übersichtlich sind die mittels „*Rechteckanordnung*“ hervorgehenden einfachen Reihen gebildet. Vgl. London, *Math. Ann.* **53**, 360 (1900), wo der Satz bewiesen wird: *Zur Konvergenz einer Doppelreihe ist notwendig und hinreichend, daß sämtliche aus ihr mittels Rechteckanordnung hervorgehenden einfachen Reihen konvergieren.*

Ist eine Doppelreihe mit positiven Gliedern konvergent, so konvergiert zur gleichen Summe die nach *Diagonalen* geordnete einfache Reihe $\sum_{\lambda}(u_{1,\lambda} + u_{2,\lambda-1} + \cdots u_{\lambda,1})$.

Eine Doppelreihe heißt *absolut konvergent*, wenn die aus den absoluten Werten ihrer Glieder gebildete Doppelreihe konvergiert. Eine absolut konvergente Doppelreihe ist immer auch *unbedingt* konvergent, d. h. sie besitzt bei jeder Anordnung der Glieder dieselbe Summe. Es gilt hiervon auch die Umkehrung (Pringsheim, *Münch. Ber.* **27**, 135 (1897)).

Wenn die Doppelreihe *absolut* konvergiert, so sind *sämtliche* einfachen Reihen, in die man sie überführen kann, konvergent und umgekehrt (London, *Math. Ann.* **53**, 360 (1900)).

Wie bei den einfach unendlichen Reihen kann man durch das Prinzip der *Reihenvergleichung Konvergenzkriterien* für Doppelreihen aufstellen. Sehr allgemeiner Natur sind die von Pringsheim, *Münch. Ber.* **27**, 144 (1897), ebda. **38**, 41 (1908), gegebenen Kriterien für Doppelreihen mit positiven Gliedern.

Es seien $\frac{1}{c_{\mu\nu}} > 0$ bezw. $\frac{1}{d_{\mu\nu}} > 0$ die allgemeinen Glieder einer konvergenten bezw. divergenten Doppelreihe. Es ist dann die Doppelreihe $\sum_{\mu}\sum_{\nu} u_{\mu\nu}$

$$\textit{konvergent}, \text{ wenn } \overline{\lim_{\mu+\nu=\infty}} c_{\mu\nu} u_{\mu\nu} < \infty,$$

$$\textit{divergent}, \text{ wenn } \underline{\lim_{\mu+\nu=\infty}} d_{\mu\nu} u_{\mu\nu} > 0$$

ist. Man kann die $c_{\mu\nu}$ und $d_{\mu\nu}$ aus den Gliedern von *einfach* unendlichen Reihen gewinnen. Sei z. B. $\frac{1}{c_\mu}$ das allgemeine Glied einer konvergenten einfach unendlichen Reihe, so kann $c_{\mu\nu} = c_\mu \cdot c_\nu$ genommen werden. Für $c_\mu = a^\mu$ $(0 < a < 1)$ folgt:

Die Doppelreihe $\sum\sum u_{\mu\nu}$ *ist konvergent oder divergent, je nachdem* $\overline{\lim_{\mu+\nu=\infty}} u_{\mu\nu}^{\frac{1}{\mu+\nu}} \lesseqgtr 1$ *ist.*

Die Definition der Doppelreihen, die Begriffe der einfachen und absoluten Konvergenz und die Konvergenzkriterien lassen sich unmittelbar auf *mehrfache Reihen* ausdehnen. Eine *p-fach unendliche Reihe* hat allgemein die Form

$$\sum_{\mu_1=1}^{\infty} \sum_{\mu_2=1}^{\infty} \cdots \sum_{\mu_p=1}^{\infty} f(\mu_1, \mu_2, \ldots \mu_p)$$

und ist *konvergent* mit der Summe S, wenn die Summen

$$s(m_1, m_2, \ldots m_p) = \sum_{\mu_1=1}^{m_1} \cdots \sum_{\mu_p=1}^{m_p} f(\mu_1, \ldots \mu_p)$$

sämtlich denselben endlichen Grenzwert S besitzen, wie auch immer die $m_1, \ldots m_p$ unabhängig voneinander ins Unendliche wachsen.

Diese Definition wird nicht wesentlich modifiziert, wenn die Summationsbuchstaben $\mu_1, \ldots \mu_p$ alle ganzzahligen Werte von $-\infty$ bis $+\infty$ durchlaufen.

Von den Konvergenzkriterien sei nur die Verallgemeinerung des Cauchyschen *Integralkriteriums* (vgl. S. 427) erwähnt:

Sei $f(x_1, x_2, \ldots x_p)$ eine immer positive Funktion von p Veränderlichen, die außerhalb einer geschlossenen Fläche S des p-dimensionalen Raumes mit wachsenden absoluten Werten der x beständig bis zu Null abnimmt, so konvergiert oder divergiert die p-fach unendliche Reihe $\sum_{\mu_1} \cdots \sum_{\mu_p} f(\mu_1, \ldots \mu_p)$, je nachdem das p-fache Integral $\int\int \cdots \int f(x_1, \ldots x_p) dx_1 \ldots dx_p$, erstreckt über den ganzen Raum außerhalb der Fläche S, einen Sinn hat oder nicht (vgl. Picard, *Traité d'analyse* **1**, 267; Riemann, *Werke*, 2. Aufl., 483).

Führt man an Stelle der Summationsbuchstaben $\mu_1, \ldots \mu_p$ neue Summationsbuchstaben mit Hilfe von linearen Substitutionen ein, so erhält man sehr allgemeine Umformungssätze für absolut konvergente mehrfache Reihen, die besonders für die Theorie der höheren Thetareihen von Wichtigkeit sind (vgl. Krazer, *Lehrbuch der Thetafunktionen* 1903, Kap. II).

§ 5. Unendliche Produkte.

Mit einer einfachen Folge von unendlich vielen Zahlen $a_1, a_2, a_3 \ldots$ bildet man der Reihe nach die Produkte

$$P_n = a_1 a_2 \ldots a_n \qquad n = 1, 2, 3, \ldots.$$

Wenn diese bei wachsendem n einem endlichen *von Null verschiedenen* Grenzwert P zustreben, so nennt man das unendliche Produkt $a_1 a_2 a_3 \ldots$ *konvergent* und P seinen *Wert*.

In allen anderen Fällen, also wenn $\lim P_n$ null oder unendlich ist oder wenn kein bestimmter Grenzwert existiert, heißt das unendliche Produkt *divergent*.

Produkte mit $\lim P_n = 0$ als divergent zu bezeichnen, ist deswegen nötig, weil sonst ein Produkt verschwinden könnte, ohne daß ein Faktor Null ist.

Für die Konvergenz eines unendlichen Produktes ist notwendig und hinreichend, daß für genügend großes n und jedes $p = 1, 2, 3 \ldots$ das Produkt $a_{n+1} a_{n+2} \ldots a_{n+p}$ der Einheit beliebig nahe gebracht werden kann.

Da bei einem konvergenten Produkt notwendig $\lim a_n = 1$ sein muß, schreibt man zweckmäßig $a_n = 1 + u_n$ $(n = 1, 2, 3 \ldots)$.

Das unendliche Produkt $\prod_1^\infty (1 + u_n)$ heißt *unbedingt* konvergent, wenn es bei jeder Anordnung seiner Faktoren denselben Wert besitzt. *Dazu ist notwendig und hinreichend, daß die Reihe* $u_1 + u_2 + u_3 + \cdots$ *unbedingt konvergiert.*

Ein konvergentes Produkt, bei dem von einem bestimmten endlichen n ab alle u_n dasselbe Vorzeichen haben, konvergiert *unbedingt*.

Ein unbedingt konvergentes Produkt konvergiert auch *absolut*, d. h. bei unbedingt konvergentem $\prod_1^\infty (1 + u_n)$ ist auch das Produkt $\prod_1^\infty (1 + |u_n|)$ konvergent. Umgekehrt zieht die absolute Konvergenz die unbedingte nach sich.

Wenn die Reihe $u_1 + u_2 + \cdots$ nur *bedingt* konvergiert, so kann das Produkt $\prod_1^\infty (1 + u_n)$ entweder bedingt konvergieren oder gegen Null divergieren, je nachdem die Reihe $u_1^2 + u_2^2 + u_3^2 + \cdots$ konvergiert oder divergiert.

Ein konvergentes Produkt läßt sich in eine Reihe:

$$\prod_1^\infty (1 + u_n) = 1 + u_1 + P_1 u_2 + P_2 u_3 + \cdots$$

verwandeln, worin $P_k = \prod_{n=1}^{k}(1 + u_n)$. Hieraus entspringt bei unbedingter Konvergenz des unendlichen Produktes die ebenfalls unbedingt konvergente Reihe

$$\prod_{1}^{\infty}(1 + u_n) = 1 + \sum u_h + \sum u_h u_i + \sum u_h u_i u_k + \cdots,$$

wobei die Summen über sämtliche Kombinationen der u_n zu je 1, 2, 3 ... zu erstrecken sind (Euler, *Introductio Cap.* 15 u. 16).

Sind die Zahlen $u_1, u_2, \ldots$ *Funktionen einer Veränderlichen* x, so wird durch das Produkt $\prod_{1}^{\infty}(1 + u_n(x))$ für jeden Wert von x, bei dem es konvergiert, eine Funktion $f(x)$ definiert.

Das Produkt $\prod(1 + u_n(x))$ heißt *gleichmäßig konvergent* in einem gewissen Bereich der Veränderlichen, wenn zu jeder positiven Zahl ε eine bestimmte *von x unabhängige* Zahl N gefunden werden kann, so daß für jedes $n > N$ und für $p = 1, 2, 3 \ldots$

$$\left| \frac{P_{n+p}}{P_n} - 1 \right| < \varepsilon$$

wird.

Sind die Funktionen $u_n(x)$ in einem Bereich stetig und konvergiert daselbst das unendliche Produkt $\prod(1 + u_n(x))$ gleichmäßig, so stellt es eine für alle Werte x des Bereichs *stetige* Funktion von x dar.

Das Produkt $\prod(1 + u_n(x))$ *konvergiert gleichmäßig, sobald die unendliche Reihe* $\sum|u_n(x)|$ *gleichmäßig konvergiert.* Alsdann konvergiert das Produkt auch *absolut.*

Die unendlichen Produkte wurden nach vereinzeltem Auftreten bei Vieta, Wallis, Dan. Bernoulli zuerst von Euler in ausgedehntem Maße benutzt. Er fand die Produktentwicklungen der trigonometrischen Funktionen, der Gammafunktion u. a. In diesem Abschnitt seien nur die ersten erwähnt:

$$\sin \pi x = \pi x \prod_{1}^{\infty}\left(1 - \frac{x^2}{n^2}\right), \qquad \cos \pi x = \prod_{1}^{\infty}\left(1 - \frac{x^2}{(n - \frac{1}{2})^2}\right),$$

die für jeden Wert von x unbedingt konvergieren. Das erste Produkt liefert für $x = \frac{1}{2}$:

$$\frac{2}{\pi} = \left(1 - \frac{1}{2^2}\right)\left(1 - \frac{1}{4^2}\right)\left(1 - \frac{1}{6^2}\right)\cdots$$

und

$$\frac{\pi}{2} = \left(1 + \frac{1}{1\cdot 3}\right)\left(1 + \frac{1}{3\cdot 5}\right)\left(1 + \frac{1}{5\cdot 7}\right)\cdots,$$

was mit dem (bedingt konvergenten) Produkt von Wallis, *Arithm. infin.* 1656:

$$\frac{\pi}{2} = \frac{2}{1}\cdot\frac{2}{3}\cdot\frac{4}{3}\cdot\frac{4}{5}\cdot\frac{6}{5}\cdot\frac{6}{7}\cdots$$

übereinstimmt. Aus diesem folgt

$$\frac{\pi}{4} = \left(1 - \frac{1}{3^2}\right)\left(1 - \frac{1}{5^2}\right)\left(1 - \frac{1}{7^2}\right)\cdots,$$

welches wieder unbedingt konvergiert, so daß sich auch

$$\left(1 - \frac{1}{2^2}\right)\left(1 - \frac{1}{3^2}\right)\left(1 - \frac{1}{4^2}\right)\cdots = \frac{1}{2}$$

ergibt.

Allgemeine Sätze über unendliche Produkte hat erst Cauchy, *Analyse algébr.* 1821 mit Hilfe des Übergangs zu Logarithmen aufgestellt. Die direkte Behandlung geht auf die Abhandlung von Weierstraß über die Theorie der analytischen Fakultäten *J. f. Math* **51**, 1 (1856) zurück. Vgl. Pringsheim, *Math. Ann.* **33**, 119 (1889) und die ausführliche Darstellung bei Stolz-Gmeiner, *Einl. in die Funktionentheorie* 1905, S. 400.

§ 6. Allgemeine Kettenbrüche.

Ein Ausdruck von der Form

$$b_0 + \frac{a_1}{b_1} + \frac{a_2}{b_2} + \frac{a_3}{b_3} + \cdots + \frac{a_n}{b_n}$$

heißt ein *Kettenbruch*. Es sind verschiedene einfachere Schreibweisen gebräuchlich; wir schreiben

$$\begin{pmatrix} a_1, & a_2, \ldots a_n \\ b_0, & b_1, & b_2, \ldots b_n \end{pmatrix}.$$

Die Größen a_i, b_i werden die *Teilzähler* und *Teilnenner*, gemeinsam auch die *Elemente*, $\frac{a_i}{b_i}$ die *Partialbrüche* des Kettenbruchs genannt.

Schließt man mit dem Teilnenner b_{k-1} ab, so heißt der so erhaltene endliche Bruch der k^{te} *Näherungsbruch* und wird mit $\frac{A_k}{B_k}$ bezeichnet.

Zähler und Nenner der Näherungsbrüche genügen den folgenden rekurrenten Beziehungen (Wallis, *Arith. infin.* 1656):

$$A_{k+1} = b_k A_k + a_k A_{k-1}$$
$$B_{k+1} = b_k B_k + a_k B_{k-1}$$

mit den Anfangswerten

$$A_0 = 1, \quad B_0 = 0$$
$$A_1 = b_0, \quad B_1 = 1.$$

A_{k+1} und B_{k+1} lassen sich durch Determinanten (*Kontinuanten*) ausdrücken (vgl. S. 69).

Die Differenz von zwei aufeinanderfolgenden Näherungsbrüchen ist

$$\frac{A_n}{B_n} - \frac{A_{n-1}}{B_{n-1}} = (-1)^n \frac{a_1 a_2 \cdots a_{n-1}}{B_{n-1} B_n},$$

und es läßt sich danach jeder Näherungsbruch durch eine *Reihe*

$$\frac{A_n}{B_n} = b_0 + \frac{a_1}{B_1 B_2} - \frac{a_1 a_2}{B_2 B_3} + \frac{a_1 a_2 a_3}{B_3 B_4} - \cdots + (-1)^n \frac{a_1 a_2 \cdots a_{n-1}}{B_{n-1} B_n}$$

ausdrücken.

Wenn alle a und b positiv sind, so sind die Differenzen zwischen zwei aufeinanderfolgenden Näherungsbrüchen abwechselnd positiv und negativ und nehmen ihrem absoluten Wert nach ab; die Näherungsbrüche mit geradem Index bilden eine abnehmende, die mit ungeradem Index eine zunehmende Reihe, und ein Näherungsbruch liegt immer zwischen den zwei vorhergehenden.

Wenn die Reihe der Teilzähler nnd Teilnenner nicht abbricht, so hat man einen *unendlichen* Kettenbruch, und wenn für $n = \infty$ ein Grenzwert für $\frac{A_n}{B_n}$ existiert, so nennt man den Kettenbruch *konvergent*. Es besteht dann ein System von unendlich vielen linearen Gleichungen von der Form

$$x = b_0 + \frac{a_1}{x_1}, \quad x_1 = b_1 + \frac{a_2}{x_2}, \ldots.$$

Es folgt aber nicht umgekehrt aus der Existenz eines solchen Systems von Gleichungen, daß sich x durch den Ketten-

bruch $\left(\begin{matrix} a_1, & a_2, \dots \\ b_0, & b_1, \dots \end{matrix}\right)$ darstellen lasse, vielmehr müssen dazu die Zahlen a_k, b_k, x_k noch gewisse Bedingungen erfüllen, z. B. daß die $|a_k|$ unter einer von k unabhängigen Schranke bleiben und $\lim_{k=\infty} \frac{1}{x_k} = 0$ ist. Perron, *Münch. Ber.* **37**, 495 (1907). Ein unendlicher Kettenbruch läßt sich in die unendliche Reihe

$$x = b_0 + \frac{a_1}{B_1 B_2} - \frac{a_1 a_2}{B_2 B_3} + \frac{a_1 a_2 a_3}{B_3 B_4} \cdots$$

verwandeln (Euler, *Introductio* Cap. 18) und konvergiert oder divergiert gleichzeitig mit dieser Reihe.

Es gibt konvergente Kettenbrüche, die durch Weglassung einer endlichen Anzahl von Anfangselementen divergent werden; sie heißen nach Pringsheim, *Münch. Ber.* **28**, 299 (1898) *bedingt konvergent*. Bleibt die Konvergenz bei solchen Weglassungen erhalten, so heißt der Kettenbruch *unbedingt konvergent*. Pringsheim hat für Kettenbrüche mit ganz beliebigen (auch *komplexen*) Elementen das folgende hinreichende *Konvergenzkriterium* aufgestellt:

Der Kettenbruch $\left(\begin{matrix} a_1, & a_2, & a_3 \dots \\ b_1, & b_2, & b_3 \dots \end{matrix}\right)$ *ist unbedingt konvergent, wenn* $|b_1| \geqq 1$ *und* $|b_k| \geqq |a_k| + 1$ *für* $k > 1$.

Nur wenn für *jedes* $k > 1$ $|b_k| = |a_k| + 1$ ist, wird dieses Kriterium etwas modifiziert (vgl. Pringsheim, *Münch. Ber.* **35**, 364 (1905)).

Sind alle Elemente *positiv*, so ist zur *Konvergenz* des Kettenbruchs *notwendig und hinreichend*, daß wenigstens eine der Reihen

$$\sum_1^\infty \frac{a_1 a_3 \cdots a_{2k-1}}{a_2 a_4 \cdots a_{2k}} b_{2k} \quad \text{und} \quad \sum_1^\infty \frac{a_2 a_4 \cdots a_{2k}}{a_1 a_3 \cdots a_{2k-1}} \frac{b_{2k+1}}{a_{2k+1}}$$

divergiert. Sind aber beide Reihen konvergent, so oszilliert der Kettenbruch zwischen zwei endlichen Grenzen. Seidel, *Unters. über die Konvergenz und Divergenz der K.*, München 1846; Stern, *J. f. Math.* **37**, 269 (1848).

Einfacher anzuwenden sind folgende Kriterien, die aber nur hinreichende, nicht notwendige Bedingungen für die Konvergenz geben:

Ein Kettenbruch mit nur positiven Elementen *konvergiert*, wenn die Reihe $\sum \frac{b_k b_{k+1}}{a_{k+1}}$ oder auch schon $\sum \sqrt{\frac{b_k b_{k+1}}{a_{k+1}}}$ *di-*

vergiert. Ebenso *konvergiert* ein derartiger Kettenbruch, wenn von einem bestimmten n ab $b_n \geqq a_n$ ist und die Reihe

$$b_1 + b_2 + b_3 + \cdots$$

divergiert.

Direkte Konvergenzkriterien für Kettenbrüche mit positiven Elementen ohne Vermittlung von Reihen hat Perron, *Münch. Ber.* **35**, 315 (1905) gegeben.

Jeder Kettenbruch mit ganzzahligen Elementen, bei dem $0 < a_k \leqq b_k$ *ist, konvergiert.* Sein Wert ist immer *irrational* und liegt zwischen b_0 und $b_0 + 1$ (Legendre, *Éléments de Géométrie*; Stern, *J. f. Math.* **11**, 38 (1834).

Der Kettenbruch

$$b_0 - \frac{1}{b_1} - \frac{1}{b_2} - \cdots$$

konvergiert, wenn von einem bestimmten n ab stets $b_n \geqq 2$ *ist.*

Die Hauptformel zur *Verwandlung einer Reihe in einen Kettenbruch* ist von Euler, *Introd.* Cap. 18, § 368 aufgestellt worden:

$$\sum_{\nu=1}^{\infty} (-1)^{\nu-1} c_\nu = \begin{pmatrix} c_1, & c_2, & c_1 c_3, & c_2 c_4, \cdots \\ 1, & c_1 - c_2, & c_2 - c_3, & c_3 - c_4, \cdots \end{pmatrix},$$

wobei Reihe und Kettenbruch gleichzeitig konvergieren. Sie führt zu Kettenbruchentwicklungen für die wichtigsten transzendenten Zahlen und Potenzreihen, von denen die folgenden mitgeteilt seien:

$$\frac{4}{\pi} = \begin{pmatrix} 1, & 9, & 25, & 49, \ldots \\ 1, & 2, & 2, & 2, \ldots \end{pmatrix} \qquad \text{(Brouncker 1659).}$$

$$\frac{1}{ln\,2} = \begin{pmatrix} 1, & 4, & 9, & 16, \ldots \\ 1, & 1, & 1, & 1, \ldots \end{pmatrix}$$

$$\frac{1}{e-1} = \begin{pmatrix} 1, & 2, & 3, & 4, & 5, \ldots \\ 1, & 2, & 3, & 4, & 5, \ldots \end{pmatrix}$$

$$= \begin{pmatrix} 1, & -1, & 1, & -1, & 1, & -1, \ldots \\ 2, & 3, & 2, & 5, & 2, & 7, \ldots \end{pmatrix}$$

(Minkowski, *Math. Ann.* **54**, 118 (1901))

$$x\frac{e^x-1}{e^x+1} = \begin{pmatrix} x^2, & x^2, & x^2, & x^2, \dots \\ 2, & 6, & 10, & 14, \dots \end{pmatrix}$$

$$-x\operatorname{tg} x = \begin{pmatrix} -x^2, & -x^2, & -x^2, & -x^2, \dots \\ 1, & 3, & 5, & 7, \dots \end{pmatrix}.$$

Die beiden letzten Entwicklungen, die bereits Euler gefunden aber nicht streng bewiesen hatte, sind in völlig einwandfreier Weise unter Nachweis der Konvergenz von Lambert, *Berl. Akad.* 1761, S. 268 durch ein dem Euklidischen Algorithmus nachgebildetes Divisionsverfahren abgeleitet worden (vgl. Pringsheim, *Münch. Ber.* **28**, 331 (1898)). Er hat mit ihnen die *Irrationalität* von e^x und $\log x$, $\operatorname{tg} x$ und $\operatorname{arctg} x$ für jedes rationale x, *also auch von der Zahl* π bewiesen.

Über Kettenbruchentwicklungen von Funktionen vgl. Padé, *Ann. éc. Norm.* (3) **9**, 7 (1892), ebda (3) **16**, 395 (1899), Stieltjes, *Toulouse Ann.* **8** (1894), **9** (1895). Über die Entwicklung des *Quotienten von zwei Potenzreihen*, wofür die von Gauß für den Quotienten zweier hypergeometrischer Reihen gefundene Entwicklung ein wichtiges Beispiel ist, vgl. Muir, *Edinb. Trans.* **27**, 467 (1876).

Ein Kettenbruch heißt *periodisch*, wenn sich seine Partialbrüche von einer bestimmten Stelle an in derselben Ordnung wiederholen, *rein periodisch*, wenn die Periode beim ersten Element beginnt, *gemischt periodisch*, wenn ihr Elemente vorangehen.

Sobald ein periodischer Kettenbruch konvergiert, ist sein Grenzwert Lösung einer quadratischen Gleichung, die aus den beiden letzten Näherungsbrüchen bis zum Schluß der ersten Periode zu bilden ist (näheres § 7).

Sind alle Elemente a_k, b_k *reell* und *positiv*, so konvergiert der periodische Kettenbruch.

Als einfachste Beispiele von periodischen Kettenbrüchen seien erwähnt:

$$\sqrt{a^2+b} = \begin{pmatrix} b, & b, \dots \\ a, & 2a, & 2a, \dots \end{pmatrix}$$ Bombelli, *Algebra* 1572.

Die beiden Lösungen der Gleichung $x^2 - bx - a = 0$ sind (Cataldi, 1613).

$$x_1 = \begin{pmatrix} a, & a, \dots \\ b, & b, & b, \dots \end{pmatrix}, \quad x_2 = -\begin{pmatrix} a, & a, & a, \dots \\ b, & b, & b, \dots \end{pmatrix}.$$

§ 7. Regelmäßige Kettenbrüche.

Der Kettenbruch $\begin{pmatrix} a_1, & a_2, & a_3, \dots \\ b_0, & b_1, & b_2, & b_3, \dots \end{pmatrix}$ läßt sich durch den völlig gleichwertigen $\begin{pmatrix} c_1 a_1, & c_1 c_2 a_2, & c_2 c_3 a_3, \dots \\ b_0, & c_1 b_1, & c_2 b_2, & c_3 b_3, \dots \end{pmatrix}$ ersetzen mit beliebigen Zahlen $c_1, c_2, c_3 \dots$ Durch geeignete Wahl der c_ν kann man sämtlichen Teilzählern oder Teilnennern bestimmte Werte vorschreiben, insbesondere kann man erreichen, daß *alle Teilzähler gleich* 1 werden. Man erhält dann die *Hauptform*:

$$q_0 + \frac{1}{q_1} + \frac{1}{q_2} + \cdots$$

in abgekürzter Schreibweise $(q_0, q_1, q_2, \dots)$. Sind hierin noch die $q_0, q_1, q_2, \dots$ *ganze positive Zahlen*, so hat man einen *regelmäßigen Kettenbruch*. Die regelmäßigen Kettenbrüche bilden die wichtigste Gattung der Kettenbrüche. Von ihnen ist im folgenden ausschließlich die Rede.

Jede positive und — wenn man negative Werte für q_0 zuläßt — *negative reelle Zahl läßt sich durch einen regelmäßigen Kettenbruch darstellen.*

Die Kettenbruchentwicklung einer *rationalen* Zahl wird mit Hilfe des *Euklidischen Algorithmus* zur Aufsuchung des *größten gemeinsamen Teilers* gefunden. Man erhält stets einen *endlichen* Kettenbruch.

Für eine *irrationale* Zahl x findet man einen *unendlichen*, stets konvergenten Kettenbruch $(q_0, q_1, q_2, \dots)$ mit Hilfe der Kette von Gleichungen:

$$\begin{aligned} x &= q_0 + \frac{1}{x_1} \\ x_1 &= q_1 + \frac{1}{x_2} \\ x_2 &= q_2 + \frac{1}{x_3} \\ &\vdots \end{aligned},$$

wobei die $q_0, q_1, q_2, \dots$ die größten[1]) in $x, x_1, x_2, \dots$ enthaltenen ganzen Zahlen bedeuten. *Umgekehrt hat ein unendlicher Kettenbruch stets einen irrationalen Wert.*

---

1) Über Kettenbrüche nach *nächsten* ganzen Zahlen vgl. Hurwitz *Acta math.* **12**, 367, (1889).

Die *Näherungsbrüche* des Kettenbruchs sind

$$\frac{P_\nu}{Q_\nu} = (q_0, q_1, q_2, \cdots q_{\nu-1}) \qquad \nu = 1, 2, 3, \ldots$$

oder durch *Gaußsche Klammern* ausgedrückt (Gauß, *Disqu. arithm.* Art. 27, jedoch schon bei Euler, *Comm. Acad. Petrop.* **7**, (1734/35), 46):

$$\frac{P_\nu}{Q_\nu} = \frac{[q_0, q_1, q_2, \cdots q_{\nu-1}]}{[q_1, q_2, \cdots q_{\nu-1}]}.$$

Zur Berechnung der Näherungsbrüche dienen die Rekursionsformeln:

$$P_{\nu+1} = q_\nu P_\nu + P_{\nu-1}, \quad Q_{\nu+1} = q_\nu Q_\nu + Q_{\nu-1}.$$
$$P_0 = 1, \qquad Q_0 = 0.$$
$$P_1 = q_0, \qquad Q_1 = 1.$$
$$P_\nu Q_{\nu-1} - Q_\nu P_{\nu-1} = (-1)^\nu.$$

Alle Näherungsbrüche sind irreduzibel; sie sind abwechselnd größer und kleiner als der Wert x des Kettenbruchs und nähern sich ihm immer mehr in der Weise, daß

$$\left| x - \frac{P_\nu}{Q_\nu} \right| < \frac{1}{Q_\nu^2}.$$

Jeder Bruch $\frac{P}{Q}$, *für den* $\left| x - \frac{P}{Q} \right| < \frac{1}{2Q^2}$ *ist, ist notwendig einer der Näherungsbrüche in der Kettenbruchentwicklung von x*, aber nicht jeder Näherungsbruch hat diese Eigenschaft, jedoch von zwei aufeinanderfolgenden Näherungsbrüchen immer mindestens einer (Vahlen, *J. f. Math.* **115**, 223 (1895)). *Es läßt sich aber eine besondere Art der Kettenbruchentwicklung angeben, bei der nur diese Näherungsbrüche auftreten* (Minkowski, *Math. Ann.* **54**, 91 (1901)). Dabei sind alle Teilzähler ± 1, alle Teilnenner positive ganze Zahlen.

Über *Nebennäherungsbrüche* eines regelmäßigen Kettenbruchs vgl. Serret, *Algebra* übers. von Wertheim, 2. Aufl. **1**, 18.

Es ist

$$\frac{P_\nu}{P_{\nu-1}} = (q_{\nu-1}, q_{\nu-2}, \cdots q_2, q_1, q_0)$$
$$\frac{Q_\nu}{Q_{\nu-1}} = (q_{\nu-1}, q_{\nu-2}, \cdots q_2, q_1);$$

Ein endlicher Kettenbruch, bei dem die von den Enden gleich weit abstehenden Teilnenner einander gleich sind, heißt

symmetrisch. Er kann eine gerade oder ungerade Anzahl von Teilnennern besitzen.

Wenn $Q^2 \pm 1$ durch P teilbar und $P > Q$ ist, so liefert die Entwicklung von $\frac{P}{Q}$ einen symmetrischen Kettenbruch, und zwar mit einer *geraden* oder *ungeraden* Anzahl von Teilnennern, je nachdem das *obere* oder *untere* Zeichen gilt.

Bei einem symmetrischen Kettenbruch mit einer geraden Anzahl von Teilnennern

$$\frac{P_{2n}}{Q_{2n}} = (q_0, q_1, \cdots q_{n-1}, q_{n-1}, \cdots q_1, q_0)$$

bestehen die Relationen

$$P_{2n} = P_n^2 + P_{n-1}^2; \quad Q_{2n-1} = Q_n^2 + Q_{n-1}^2$$
$$P_{2n-1} = Q_{2n} = P_n Q_n + P_{n-1} Q_{n-1}$$
$$P_{2n} \cdot Q_{2n-1} = Q_{2n}^2 + 1.$$

Bei einem symmetrischen Kettenbruch mit einer ungeraden Anzahl von Teilnennern

$$\frac{P_{2n+1}}{Q_{2n+1}} = (q_0, q_1, \cdots q_{n-1}, q_n, q_{n-1}, \cdots q_1, q_0)$$

ist

$$P_{2n+1} = (P_{n+1} + P_{n-1}) P_n, \quad Q_{2n} = (Q_{n+1} + Q_{n-1}) Q_n$$
$$P_{2n} = Q_{2n+1} = P_n Q_{n+1} + Q_n P_{n-1} = Q_n P_{n+1} + P_n Q_{n-1}$$
$$P_{2n+1} \cdot Q_{2n} = Q_{2n+1}^2 - 1.$$

Zwei irrationale Zahlen x, y heißen *äquivalent*, wenn sich vier ganze Zahlen α, β, γ, δ mit der Determinante

$$\alpha\delta - \beta\gamma = \pm 1$$

bestimmen lassen, so daß

$$y = \frac{\alpha x + \beta}{\gamma x + \delta}$$

ist.

Die Kettenbruchentwicklungen von zwei äquivalenten Zahlen stimmen von einem gewissen Teilnenner an überein.

Umgekehrt ist *die notwendige und hinreichende Bedingung für die Äquivalenz von zwei irrationalen Zahlen, daß ihre Kettenbruchentwicklungen von einem gewissen Teilnenner an übereinstimmen.*

Jede reelle quadratische Irrationalzahl $\frac{A+B\sqrt{D}}{C}$ *mit ganzzahligen A, B, C und positivem ganzem D oder — anders ausgedrückt — jede Lösung einer quadratischen Gleichung mit ganzzahligen Koeffizienten, deren Diskriminante positiv und kein Quadrat ist, läßt sich durch einen periodischen Kettenbruch darstellen.* Umgekehrt ist *jeder periodische Kettenbruch Lösung einer quadratischen Gleichung mit ganzzahligen Koeffizienten.* Insbesondere ist der *rein periodische* Kettenbruch

$$x = (\overline{q_0, q_1, q_2, \cdots q_{n-1}}, \cdots)$$

die positive Lösung der Gleichung

$$Q_n x^2 + (Q_{n-1} - P_n)x - P_{n-1} = 0.$$

Die Kettenbruchentwicklung für die zweite Lösung dieser Gleichung hat die *umgekehrte Periode.* Ebenso bei gemischt periodischen Kettenbrüchen (Galois, *Ann. math.* **19**, 294 (1828)).

Bei der *Kettenbruchentwicklung der Quadratwurzel einer rationalen Zahl* geht der Periode *ein* Glied voraus; die Periode schließt mit dem *Doppelten dieses Anfangsgliedes*, und *die übrigen Teilnenner der Periode bilden eine symmetrische Reihe.* Umgekehrt stellt ein derartiger Kettenbruch stets die Quadratwurzel aus einer rationalen Zahl dar. Damit diese ganz ist, müssen die Teilnenner q_i noch weitere Bedingungen erfüllen. Vgl. Muir, *Edinb. Proceed.* 8, 229 (1874).

Besitzt in der Entwicklung von $\sqrt{D}$ *die Periode n Glieder, so besteht zwischen Zähler und Nenner der jeweils bis zum vorletzten Glied einer jeden Periode genommenen Näherungsbrüche die Beziehung*

$$P_{\mu n}^2 - DQ_{\mu n}^2 = (-1)^{\mu n} \qquad \mu = 1, 2, 3, \ldots.$$

Eine Tabelle der Kettenbruchentwicklungen für die Quadratwurzeln der unter 120 liegenden ganzen Zahlen findet sich bei Euler, *Comment. arith. coll.* 1, p. 322, abgedruckt in Wertheim, *Zahlenlehre*, 1902, S. 223. Eine Tabelle bis zu $\sqrt{1000}$ enthält Degen, *Canon Pellianus*, *Hafniae 1817.*

Als Begründer der Lehre von den Kettenbrüchen ist Bombelli (1572) anzusehen, der sie zur Berechnung von Quadratwurzeln benutzte (vgl. Wertheim, *Abh. zur Gesch. der Math.* 8, 149 (1898)). Die gleiche Methode, jedoch mit einer der heutigen Schreibweise verwandten Bezeichnung und Untersuchung der Eigenschaften der Näherungsbrüche bei Cataldi 1613. In

systematischer Weise wurden die Kettenbrüche jedoch erst von Euler behandelt (*Petrop. Comment.* **9** (1737), **11** (1739); *Introd.*; ferner *Nov. Comment. Petrop.* **9** (1764), **11** (1765)); Lagrange, Zusätze zu Eulers *Algebra* (Regelmäßige periodische K.); Legendre, *Théorie des nombres.* — *Geometrische Betrachtungen* hat in systematischer Weise F. Klein (*Vorl. üb. Zahlentheorie*) herangezogen. Eine zusammenfassende Darstellung der gesamten Kettenbruchlehre von Perron, *Die Lehre von den Kettenbrüchen,* erscheint demnächst bei B. G. Teubner, Leipzig.

Über die Beziehungen zwischen *Kettenbrüchen* und *divergenten Reihen* vgl. Padé, *Ann. de l'éc. Norm.* (3) **9**, (1892), Supplément; *Acta math.* **18**, 97 (1894); Stieltjes, *Ann. de la Fac. d. sc. de Toulouse* **8** u. **9** (1894/5); Borel, *Leç. s. l. sér. div.* 1901.

Eine auf Jacobi, *J. f. Math.* **69**, 29 (1868) zurückgehende Verallgemeinerung des Kettenbruchalgorithmus ist am eingehendsten von Perron, *Math. Ann.* **64**, 1 (1907), *Münch. Ber.* **37**, 401 (1907) untersucht worden.

Kapitel VII.

Differentialrechnung.

Von *Paul Epstein* in Straßburg i. E.

Einleitung.

Seit den ersten Jahrzehnten des 17. Jahrhunderts vollzieht sich die Erweiterung des mathematischen Gedankenkreises (*infinitesimale Betrachtungen, veränderliche Größen*), die schließlich in der Schöpfung der Differential- und Integralrechnung gipfeln sollte. Es waren in erster Linie *geometrische* Probleme, an denen sich die Fruchtbarkeit der infinitesimalen Betrachtungen bewährte: zunächst — bewußt anknüpfend an *Archimedes* — *Flächen-* und *Körperberechnungen*, dann das *direkte* und *inverse Tangentenproblem, Maxima* und *Minima, Rektifikationen*. Daneben entwickelt sich, vielfach durch *mechanische* Vorstellungen beeinflußt, der Begriff (oder besser das Gefühl) der *stetigen Veränderung*. Die folgende Übersicht mag die angedeutete Entwicklung in ihren Hauptzügen kennzeichnen. Wegen aller Einzelheiten, insbesondere wegen des Zusammenhangs mit der Philosophie des Mittelalters (Bradwardinus, Oresme, Nicolaus von Cues) muß auf Cantor, *Gesch. d. Math.* **2** und **3** verwiesen werden. Vgl. ferner Zeuthen, *Gesch. d. Math. im 16. u. 17. Jahrh.*

Kepler (1571—1630), *Stereometria doliorum* 1615, Flächen- und Körperberechnungen, Maxima und Minima. Cavalieri, *Geometria indivisibilibus continuorum . . . promota* 1635, *Exercitationes geometricae*. Flächen- und Körperberechnungen, Schwerpunktsbestimmungen. Descartes (1596—1650), *Geometrie* 1637. *Briefwechsel*, Tangentenproblem (vgl. Tannery, 3. intern. Math.-Kongr. 502 (1904)), Rektifikationen, Kurvenuntersuchungen. Fermat (1601—1665, der bedeutendste Vorgänger von Newton und Leibniz), Maxima und Minima (zwischen 1629 und 1638, veröffentl. 1679 in *Varia Opera*, enthält den Begriff des Differentialquotienten), Tangentenproblem (veröffentl. 1642), Flächen-

berechnungen (*Varia Opera* 44), Rektifikationen (*de linearum curvarum cum lineis rectis comparatione* 1660). Toricelli, *Opera Geometrica* 1644. Roberval, *Observations sur la composition des mouvements* (1668, aber schon 1644 erwähnt), Anwendung des Kräfteparallelogramms auf das Tangentenproblem. Wallis, *Arithmetica infinitorum* 1655, Potenzsummen, Grenzübergang. Pascal (1623—1662), *Œuvres* 3 (1872) Potenzsummen, Partielle Integration, Charakteristisches Dreieck. Huygens (1629—1695), *Horologium oscillatorium* 1673 (vollendet 1665). Evoluten und Evolventen. Krümmungsmittelpunkt. Maxima und Minima, Tangentenproblem (zwischen 1652 und 1667, sehr nahe übereinstimmend mit *de Sluse* 1652). Barrow (vgl. Zeuthen, 1. intern. Math. Kongr. 274 (1897)), *Lectiones opticae et geometricae* 1669. Tangentenproblem auf Grund der Bewegungslehre. Newton (1643—1727), *Analysis per aequationes* 1666. Flächen- und Körperberechnungen und Rektifikationen als gleichartige Aufgaben erkannt. *Methodus fluxionum* (1671, veröffentl. 1736). Die Bezeichnungen *Fluxion* (Differentialquotient) und *Fluente* zum ersten Mal in einem *Brief an Leibniz* vom 24. 10. 1676. Leibniz (1646—1716) (Leibniz' math. Schriften, herausg. von Gerhardt 1849—1863, Gerhardt, Die Entdeckung d. höh. Analysis 1855). Briefwechsel und Manuskripte von 1670—1675. Ausbildung der *Bezeichnung. Endgültige Lösung des Tangentenproblems* in der Antwort auf Newtons Brief.

Der große Fortschritt von Newton und Leibniz besteht in der Schöpfung eines *Algorithmus,* der die große Menge der vorausgehenden Einzeluntersuchungen unter einheitlichen Gesichtspunkten zusammenfaßt. Wem von beiden das Hauptverdienst zukommt, kann dahingestellt bleiben; sicher war Newton vor Leibniz im Besitz wichtiger Erkenntnisse, aber unbestreitbar wäre die glänzende Entwicklung der Mathematik im 18. Jahrhundert ohne die von Leibniz erfundene *Zeichensprache* nicht möglich gewesen.

Die Literatur über Differential- und Integralrechnung ist außerordentlich umfangreich. Von großer historischer Bedeutung und für die Entwicklung der Infinitesimalrechnung grundlegend sind: Euler, *Institutiones calculi differentialis*, 1755; *Institutiones calculi integralis*, 2. Ausg., 4 Bde., 1792—1794. Lagrange, *Leçons sur le calcul des fonctions*, 2. éd. 1806; *Théorie des fonctions analytiques*, 2. éd. 1813. Cauchy, *Cours d'analyse* 1821; *Résumé des leçons données à l'école polyt.* 1823; *Leçons sur le calcul différentiel,* 1829.

Von neueren Lehrbüchern seien genannt: Cesàro, *Elementares Lehrbuch d. algebr. Analysis und Infinitesimalrechnung* 1904. Kowalewski, *Die klass. Probleme d. Analysis* 1910. Serret-Scheffers, *Lehrb. d. Differential- und Integralr.*, 3 Bde. Stegemann-Kiepert, *Grundriß d. Diff.- u. Integralr.*, 2 Bde. Genocchi-Peano, *Differentialr. und Grundzüge d. Integralr.* 1899. Stolz, *Grundzüge d. Different.- und Integralr.*, 3 Bde., 1893/99. Pascal, *Calcolo infinitesimale*, 3 Bde. Jordan, *Cours d'analyse*, 3 Bde., 2. éd., 1893/96. Goursat, *Cours d'analyse*, 2 Bde., 1902.

Zur ersten Einführung sind die folgenden Werke geeignet, die besonders die Bedürfnisse der angewandten Mathematik im Auge haben: Scheffers, *Lehrb. d. höheren Mathematik*, 1905. Burkhardt, *Vorl. über die Elemente der Diff.- und Integralr.*, 1907. Nernst-Schoenflies, *Einf. in die math. Behandl. der Naturw.* Perry, *Höhere Analysis für Ingenieure*, 1910. Greenhill, *Differential- and Integralcalculus*, 1896.

§ 1. Infinitesimale Größen.

Eine Veränderliche, die bei einem bestimmten Grenzprozeß den Grenzwert *Null* hat, heißt *unendlich klein* oder *infinitesimal*. Zwei unendlich kleine Größen α und β heißen *von derselben Ordnung*, wenn $\lim \frac{\alpha}{\beta}$ einen *endlichen, von Null verschiedenen Wert* besitzt. Ist aber $\lim \frac{\alpha}{\beta} = 0$, so nennt man α von *höherer Ordnung*, ist $\lim \frac{\alpha}{\beta} = \infty$, so nennt man α von *niedrigerer Ordnung* als β.

Ist β von höherer Ordnung als α und existiert eine positive Zahl n derart, daß $\lim \frac{\beta}{\alpha^n}$ einen endlichen von Null verschiedenen Wert besitzt, so sagt man, β sei *von der Ordnung n in bezug auf* α. Nicht immer läßt sich eine derartige Zahl n finden. So ist z. B., wenn $\beta = \frac{1}{\log \alpha}$ ist:

$$\lim \frac{\beta}{\alpha^n} = \begin{cases} 0 & \text{für } n \leqq 0 \\ \infty & \text{„ } n > 0. \end{cases}$$

Eine infinitesimale Größe bleibt infinitesimal von derselben Ordnung, wenn sie mit einer endlichen von Null verschiedenen Größe multipliziert wird.

Zwei infinitesimale Größen von gleicher Ordnung unter-

scheiden sich im allgemeinen um eine Infinitesimale derselben Ordnung; nur wenn ihr Verhältnis den *Grenzwert 1* besitzt, unterscheiden sie sich um eine Infinitesimale von *höherer* Ordnung.

Der Grenzwert des Verhältnisses von zwei infinitesimalen Größen ändert sich nicht, wenn man zu ihnen Infinitesimale höherer Ordnung hinzufügt.

Die algebraische Summe einer endlichen Anzahl von Infinitesimalen ist wieder infinitesimal, und ihre Ordnung ist mindestens gleich der niedrigsten in der Summe vorkommenden Ordnung.

Das Produkt von zwei infinitesimalen Größen α und β, die in bezug auf eine dritte Infinitesimale γ von den Ordnungen m und n sind, ist infinitesimal von der Ordnung $m + n$ in bezug auf γ.

§ 2. Begriff und Existenz des Differentialquotienten.

Wenn $y = f(x)$ eine Funktion der reellen Veränderlichen x ist und man dem x einen beliebigen Zuwachs $\Delta x = h$ gibt, so erfährt im allgemeinen auch y einen Zuwachs, der Δy sei.

Die Grenze des Verhältnisses $\frac{\Delta y}{\Delta x}$ für $\Delta x = 0$ oder

$$\lim_{h=0} \frac{f(x+h) - f(x)}{h}$$

heißt, wenn sie existiert und unabhängig ist von der Art und Weise, in der Δx nach Null konvergiert, die *Derivierte* (*Ableitung*) *der Funktion im Punkt x.* Sie wird durch $f'(x)$ oder y' bezeichnet.

Ein *infinitesimaler*, im übrigen aber beliebiger Zuwachs der unabhängigen Variablen x heißt *Differential* von x und wird mit dem Symbol dx bezeichnet.

Differential der Funktion y von x heißt das Produkt aus der Ableitung von y mit dem Differential der unabhängigen Variablen (Cauchy, *Exerc. d'anal. et de phys.* **3**, 7 ff. (1844)). Man bezeichnet es mit dy, und es ist also

$$dy = f'(x) \cdot dx.$$

Man schreibt deshalb auch die Derivierte

$$f'(x) = \frac{dy}{dx}$$

und nennt sie den *Differentialquotienten* der Funktion y nach x. Ist $f'(x)$ stetig, so *unterscheidet sich das Differential der Funktion um Unendlichkleine höherer Ordnung von dem Zuwachs, den die*

Funktion erfährt, wenn man der unabhängigen Variablen den Zuwachs dx gibt.

Es kann vorkommen, daß das Verhältnis $\frac{\Delta y}{\Delta x}$ verschiedene Grenzwerte besitzt, je nachdem Δx von der *positiven* oder *negativen* Seite her die Grenze *null* erreicht; alsdann spricht man von der Derivierten zur *Rechten* und der zur *Linken* von x. (*Vorwärts* und *rückwärts genommener Differentialquotient.*) Unterscheidet man noch genauer zwischen *oberer* und *unterer* Grenze (vgl. Kap. I, § 10), so hat man *vier* Derivierte zu betrachten. Nur wenn diese vier Grenzwerte einander gleich sind, nennt man die Funktion *differenzierbar*.

Allgemeine *notwendige* und *hinreichende* Kriterien für die Differenzierbarkeit einer Funktion lassen sich nicht angeben. Sie ist, wenn sie vorhanden ist, ebenso eine fundamentale Eigenschaft der Funktion wie etwa die Stetigkeit. *Notwendige* Bedingung für die Existenz einer Ableitung in einem Punkt x ist jedenfalls, daß die Funktion in diesem Punkte *stetig* und in ihm und seiner Umgebung *endlich* ist, aber es ist nicht, wie man früher glaubte, jede stetige Funktion auch differenzierbar. Ein von Ampère, *J. éc. pol.* **13**, 148 (1806) versuchter Beweis dieser Behauptung ist nicht stichhaltig (vgl. Du Bois-Reymond, *Journ. f. Math.* **79**, 28 (1875)). Das erste Beispiel einer stetigen und trotzdem in unendlich vielen Punkten *nicht differenzierbaren* Funktion hat Riemann in seiner 1854 verfaßten, aber erst 1867 veröffentlichten Habilitationsschrift gegeben. (*Über die Darstellbarkeit einer Funktion durch eine trigonometrische Reihe. Werke* 2. Aufl. S. 266.) Diese Funktion entsteht durch Integration der Funktion

$$f(x) = \sum_{n=1}^{\infty} \frac{(nx)}{n^2},$$

wobei das Zeichen (z), sobald z in der Mitte zwischen zwei ganzen Zahlen liegt, den Wert *null*, sonst aber den Überschuß von z über die *nächste* ganze Zahl bedeutet. Das von Riemann angegebene Beispiel führte Hankel zu dem sogenannten *Kondensationsprinzip*, mit dessen Hilfe man, wenn eine Funktion mit einer Singularität in einem Punkt gegeben ist, eine andere Funktion bilden kann, welche dieselbe Singularität in unendlich vielen Punkten besitzt. (Hankel, *Untersuchungen üb. die unendlich oft oszillierenden und unstetigen Funktionen*, Tübingen 1870; *Math. Ann.* **20**, 63 (1882).) Auf Grund dieses Prinzips

lassen sich Funktionen bilden, die in unendlich vielen (aber nicht in allen) Punkten eines Intervalls keine Derivierte haben. Vgl. Dini, *Theorie d. Funkt. einer reellen Größe*, Leipzig 1892, S. 157 ff. Einfacher und weiter tragend ist ein von G. Cantor (*Math. Ann.* **19**, 588 (1882)) angegebenes Kondensationsprinzip. (Dini, a. a. O. S. 188 ff.) Viele Literaturangaben bei Jourdain, *The development of the theory of transfinite numbers, Arch. f. Math. u. Phys.* (3) **10** (1906).

Das bekannteste Beispiel einer stetigen *nirgends differenzierbaren* Funktion gab Weierstraß in seinen Vorlesungen; es wurde zuerst veröffentlicht von Du Bois-Reymond, *Journ. f. Math.* **79**, 29 (1875). Vgl. Wiener, *Journ. f. Math.* **90**, 221 (1881), Weierstraß, *Funktionenlehre* 1886, S. 100, und die besonders anschauliche Darstellung bei F. Klein, *Anw. d. Diff. u. Integralr. auf Geometrie. Autographiertes Vorlesungsheft* 1901. Die Funktion lautet:

$$f(x) = \sum_{n=0}^{\infty} a^n \cos(b^n \pi x),$$

worin a zwischen 0 und 1, b eine ungrade ganze Zahl und $ab > 1 + \frac{3}{2}\pi$ ist. Sehr allgemeine Funktionen von derselben Eigenschaft, die die Weierstraßsche Funktion als speziellen Fall enthalten, hat Dini (a. a. O. S. 218 ff.) aufgestellt. Vgl. Lerch, *Journ. f. Math.* **103**, 126 (1888). Von weiteren Arbeiten über stetige nicht differenzierbare Funktionen seien genannt H. A. Schwarz, *Ges. Abhandl.* **2**, 269, Darboux, *Ann. éc. norm.* (2), **4**, 57 (1875), Steinitz, *Math. Ann.* **52**, 58 (1899), H. v. Koch, *Acta math.* **30**, 145 (1907), Landsberg, *Math. Ver.* **17**, 46 (1908), Faber, *Math. Ann.* **66**, 81 (1908). Köpcke (*Math. Ann.* **34**, 161 (1889) und **35**, 104 (1890)) hat gezeigt, daß eine Funktion in jedem endlichen Intervall unendlich viele Oszillationen haben und doch differenzierbar sein kann. Vgl. Schoenflies, *Math. Ann.* **54**, 553 (1901) und desselben Verf. *Bericht üb. die Mengenlehre, Math. Ver.* **8**, 161 ff. Man sieht daraus, daß die Eigenschaft der *Stetigkeit* und *Differenzierbarkeit* noch nicht ausreicht, um eine Funktion $y = f(x)$ als eine *gewöhnliche* (*reguläre*) Funktion zu kennzeichnen, d. h. eine solche, die sich geometrisch durch eine gewöhnliche kontinuierliche *Kurve* veranschaulichen läßt, wenn man x, y als rechtwinklige Punktkoordinaten deutet. Eine derartige Funktion muß nämlich jedenfalls auch *abteilungsweise monoton* sein, d. h. sie darf in jedem endlichen Intervall nur eine endliche Anzahl

Maxima und Minima besitzen. Diese Eigenschaft machte sich zum ersten Mal geltend in der Arbeit von Dirichlet über die Theorie der Fourierschen Reihe (*Journ. f. Math.* **4** (1829), *Werke*, **1**, 131. Vgl. Du Bois-Reymond, *Journ. f. Math.* **79**, 33 (1875)). Über die sonstigen Bedingungen, die eine reguläre Funktion erfüllen muß, vgl. etwa die erwähnte Vorlesung von F. Klein. Entsprechend dem geometrischen Ursprung des Funktionsbegriffes hat man früher (bis zur Aufstellung des allgemeinsten Dirichletschen Funktionsbegriffs (Kap. I, § 9, vgl. Hankel, *Math. Ann.* **20**, 67 (1870)) einer „stetigen" Funktion stillschweigend die Eigenschaften einer *regulären* Funktion zugeschrieben, und es soll auch im folgenden, wo nichts anderes ausdrücklich bemerkt ist, unter einer „Funktion" schlechthin immer eine reguläre Funktion verstanden werden. Wir können dann diesen Paragraphen mit dem folgenden Satz schließen: *Stellt man die Funktion* $y = f(x)$ *geometrisch durch eine Kurve dar, so gibt der Differentialquotient* $\frac{dy}{dx} = f'(x)$ *die trigonometrische Tangente des Winkels an, den die Kurventangente im Punkte* (x, y) *mit der positiven Richtung der x-Achse bildet.*

§ 3. Grundregeln für die Differentiation entwickelter Funktionen einer Veränderlichen.

1. *Der Differentialquotient einer Konstanten ist Null.*

2. *Der Differentialquotient einer algebraischen Summe von Funktionen ist gleich der algebraischen Summe der Differentialquotienten der einzelnen Funktionen.*

Eine in einem Intervall konvergente *unendliche Reihe von Funktionen*

$$F(x) = f_1(x) + f_2(x) + f_3(x) + \cdots$$

heißt *gliedweise differenzierbar,* wenn die aus den Derivierten der einzelnen Funktionen gebildete Reihe konvergiert und als Summe die Derivierte von $F(x)$ ergibt. Es gilt der Satz:

3. *Eine konvergente unendliche Reihe von Funktionen ist gliedweise differenzierbar, wenn die einzelnen Derivierten stetig sind und die aus ihnen gebildete Reihe gleichmäßig konvergiert.*

Insbesondere läßt sich eine *Potenzreihe* für jeden Wert der Variablen, bei dem sie konvergiert, beliebig oft gliedweise differenzieren.

4. *Der Differentialquotient eines Produkts aus einer Konstanten mit einer Funktion ist gleich dem Produkt der Konstanten mit dem Differentialquotienten der Funktion:*

$$\frac{d(ay)}{dx} = a\frac{dy}{dx},$$

wenn a konstant.

5. *Ein Produkt von zwei Funktionen wird differentiiert, indem man jede Funktion mit der Derivierten der andern multipliziert und die erhaltenen Produkte addiert:*

$$(uv)' = uv' + vu'.$$

6. *Den Differentialquotienten eines Produktes von mehr als zwei Funktionen findet man am besten nach der folgenden Formel (logarithmische Differentiation):*

$$\frac{(u_1 u_2 \dots u_n)'}{u_1 u_2 \dots u_n} = \frac{u_1'}{u_1} + \frac{u_2'}{u_2} + \dots + \frac{u_n'}{u_n}.$$

7. *Die Derivierte eines Quotienten von zwei Funktionen ist:*

$$\left(\frac{u}{v}\right)' = \frac{vu' - uv'}{v^2}.$$

8. Ist $x = \varphi(y)$ die *inverse Funktion* von $y = f(x)$, so ist *der Differentialquotient von x nach y der reziproke Wert des Differentialquotienten von y nach x:*

$$\varphi'(y) = \frac{1}{f'(x)}.$$

9. Sehr oft kann man eine Funktion $y = F(x)$ nur differentiieren, indem man sie als (einfachere) *Funktion* $f(z)$ *einer anderen Funktion* $z = \varphi(x)$ auffaßt, also $y = f(\varphi(x))$. Man nennt dann z *Zwischenvariable* und hat den Satz:

Man findet den Differentialquotienten von y nach x, indem man zunächst y nach z differentiiert und mit dem Differentialquotienten der Zwischenvariablen multipliziert:

$$\frac{dy}{dx} = \frac{dy}{dz} \cdot \frac{dz}{dx}.$$

10. Wir stellen hier die *elementaren Funktionen und ihre Differentialquotienten* zusammen:

$$y = x^n, \qquad y' = n x^{n-1}$$

$$y = e^x, \qquad y' = e^x$$

$$y = a^x, \qquad y' = a^x \ln a$$

$$y = \ln x, \qquad y' = \frac{1}{x}$$

$$y = \sin x, \qquad y' = \cos x$$

$$y = \cos x, \qquad y' = -\sin x$$

$$y = \operatorname{tg} x, \qquad y' = \frac{1}{\cos^2 x}$$

$$y = \operatorname{ctg} x, \qquad y' = -\frac{1}{\sin^2 x}$$

$$y = \operatorname{arc\,sin} x, \qquad y' = \frac{1}{\sqrt{1-x^2}} \qquad \begin{matrix} -1 < x < +1 \\ -\frac{\pi}{2} < y < +\frac{\pi}{2} \end{matrix}$$

$$y = \operatorname{arc\,cos} x, \qquad y' = -\frac{1}{\sqrt{1-x^2}} \qquad \begin{matrix} -1 < x < +1 \\ 0 < y < \pi \end{matrix}$$

$$y = \operatorname{arc\,tg} x, \qquad y' = \frac{1}{1+x^2}$$

$$y = \operatorname{arc\,ctg} x, \qquad y' = -\frac{1}{1+x^2}.$$

§ 4. Differentialquotienten höherer Ordnung von Funktionen einer Veränderlichen.

Wiederholt man an der Derivierten von y die Operation des Differentiierens und fährt so fort (vorausgesetzt, daß man immer differentiierbare Funktionen hat), so erhält man die *Differentialquotienten* 2ter, 3ter, ... *Ordnung* oder die *zweiten, dritten. ... Derivierten.* Man schreibt sie

$$y'' = f''(x) = \frac{d^2 y}{dx^2},$$

$$y''' = f'''(x) = \frac{d^3 y}{dx^3},$$

und allgemein

$$y^{(n)} = f^{(n)}(x) = \frac{d^n y}{dx^n}.$$

Es ist also *das n^{te} Differential der Funktion y*:

$$d^n y = f^{(n)}(x)(dx)^n,$$

unter der Voraussetzung, daß man das erste Differential dx der unabhängigen Veränderlichen x als unabhängig von dieser Veränderlichen annimmt. (Vgl. Cesàro, *Lehrbuch der algebr. Analysis und Infinitesimalrechnung,* 1904, S. 494.)

Eine Funktion, von der man Differentialquotienten beliebiger Ordnung bilden kann, heißt *unbeschränkt differentiierbar.* (Vgl. das Referat von Pringsheim über die Grundlagen der allgemeinen Funktionenlehre in der *Encykl.* **2**, 23.) Besitzt $f(x)$ an der Stelle $x = a$ Differentialquotienten bis zur n^{ten} Ordnung von endlicher Größe, so ist $f^{(n)}(a)$ der Grenzwert des Verhältnisses $\frac{\Delta^n f(a)}{h^n}$ für $\lim h = 0$, wenn unter $\Delta^n f(a)$ die n^{te} Differenz der Funktion $f(x)$ für $x = a$:

$$\Delta^n f(a) = f(a+nh) - \binom{n}{1} f(a+(n-1)h) + \binom{n}{2} f(a+(n-2)h) - \cdots + (-1)^{n-1}\binom{n}{1} f(a+h) + (-1)^n f(a)$$

verstanden wird. Der umgekehrte Satz ist nicht richtig, d. h. es kann der Grenzwert existieren, ohne daß die n^{te} Derivierte $f^{(n)}(a)$ existiert. (Vgl. Harnack, *Math. Ann.* **23**, 260 (1884), Stolz, *Grundz. d. Diff.- u. Int.-Rechnung,* Leipzig 1893, **1**, 93.)

Sind x und y *Funktionen einer unabhängigen Veränderlichen* t und werden ihre Derivierten nach t mit x', x'', ..., y', y'', ... bezeichnet, so sind die *Derivierten von y nach x:*

$$\frac{dy}{dx} = \frac{y'}{x'}, \quad \frac{d^2y}{dx^2} = \frac{x'y'' - y'x''}{x'^3},$$

$$\frac{d^3y}{dx^3} = \frac{x'(x'y''' - y'x''') - 3x''(x'y'' - y'x')}{x'^5}.$$

Schon Leibniz hat für die n^{te} *Derivierte eines Produkts von zwei Funktionen* die Formel

$$(uv)^{(n)} = u^{(n)}v + \binom{n}{1} u^{(n-1)}v' + \binom{n}{2} u^{(n-2)}v'' + \cdots + uv^{(n)}$$

gegeben.

Über die allgemeine Theorie der höheren Differentialquotienten und ihre Literatur vgl. den Artikel von Voß in der *Enzykl.* **2**, 87.

Als Beispiele für höhere Differentialquotienten mögen folgende angeführt werden:

$$\frac{d^n(x^m)}{dx^n} = m(m-1)\ldots(m-n+1)x^{m-n};$$

$$\frac{d^n \sin x}{dx^n} = \sin\left(x + \frac{n\pi}{2}\right), \quad \frac{d^n \cos x}{dx^n} = \cos\left(x + \frac{n\pi}{2}\right).$$

$$\frac{d^n}{dx^n}(e^{ax}\cos bx) = r^n e^{ax}\cos(bx + n\vartheta),$$

$$\frac{d^n}{dx^n}(e^{ax}\sin bx) = r^n e^{ax}\sin(bx + n\vartheta),$$

worin $r = \sqrt{a^2 + b^2}$ und $\vartheta = \operatorname{arc\,tg}\frac{b}{a}$ ist.

Ist $f(x) = ax^2 + 2bx + c$ eine quadratische Funktion mit der Diskriminante $\Delta = b^2 - ac$ und wird $\frac{a(ax^2+2bx+c)}{(ax+b)^2} = u$, $\frac{\Delta}{(ax+b)^2} = 1 - u = v$ gesetzt, so ist

$$\frac{d^n}{dx^n} f(x)^\mu = 2\mu(2\mu-2)\ldots(2\mu-2n+2)\frac{(ax+b)^n}{f(x)^{n-\mu}} F\left(-\frac{n}{2}, -\frac{n-1}{2}, \mu-n+1, u\right)$$

$$= 2\mu(2\mu-1)\ldots(2\mu-n+1)\frac{(ax+b)^n}{f(x)^{n-\mu}} F\left(-\frac{n}{2}, -\frac{n-1}{2}, \frac{1}{2}-\mu, v\right),$$

wobei die F abbrechende hypergeometrische Reihen in der Bezeichnung von Gauß bedeuten. Die erste Formel bei Schlömilch, *Übungsb.* 5. Aufl. Leipzig 1904, **1**, 52; die zweite für $\mu = -\frac{1}{2}$ bei Hermite, *Cours d'analyse* 1873, p. 310, für dasselbe μ und $f(x) = 1 - x^2$ schon bei Euler, *Inst. calc. diff.* 1755, p. 150.

Für $n = \varkappa - 1$, $\mu = \varkappa - \frac{1}{2}$ erhält man

$$\frac{d^{\varkappa-1}}{dx^{\varkappa-1}} f(x)^{\varkappa-\frac{1}{2}} = \frac{1\cdot 3\cdot 5\cdots 2\varkappa-1}{a^{\frac{1}{2}}\cdot\varkappa}(ax+b)^\varkappa \cdot \frac{(1+\sqrt{u})^\varkappa - (1-\sqrt{u})^\varkappa}{2},$$

und dies geht für $f(x) = 1 - x^2$ in eine Formel von Jacobi über:

$$\frac{d^{\varkappa-1}(1-x^2)^{\varkappa-\frac{1}{2}}}{dx^{\varkappa-1}} = (-1)^{\varkappa-1}\frac{1\cdot 3\cdot 5\cdots 2\varkappa-1}{\varkappa}\sin(\varkappa \arccos x).$$

Für $\mu = -\frac{1}{2}$ bzw. $\mu = n$ liefert die erste bzw. zweite der obigen Formeln:

$$\frac{d^n}{dx^n}\left(\frac{1}{\sqrt{f(x)}}\right) = (-1)^n n! \frac{a^{\frac{n}{2}}}{f(x)^{\frac{n+1}{2}}} P_n\left(\frac{1}{\sqrt{u}}\right),$$

$$\frac{d^n}{dx^n} f(x)^n = 2\cdot 4\cdot 6\cdots 2n\cdot \Delta^{\frac{n}{2}} P_n\left(\frac{1}{\sqrt{v}}\right),$$

worin P_n die (auch oft mit X_n bezeichnete) *einfache Kugelfunktion*,

auch *Legendresches Polynom* genannt, bedeutet. Zur ersten Formel vgl. Thomson u. Tait, *Natural philos.*, deutsche Ausg. 1871, **1**, 175. Die letzte Formel ist für $f(x) = x^2 - 1$, also $\Delta = 1$, $\frac{1}{\sqrt{v}} = x$ von Rodrigues, *Corr. d. l'éc. polyt.* **3** (1816), 375 gefunden worden. Über Verallgemeinerungen dieser Formel durch Jacobi, Heine, Hermite u. a. vgl. Burkhardt, *Bericht über Entwickl. nach oszill. Funktionen, Math.-Ver.* **10**, § 84 und 85.

Über die Ausdehnung des Begriffs der höheren Differentialquotienten auf nicht ganzzahlige n (*Differentiationen zu beliebigem Index*), die schon von Leibniz, Joh. Bernoulli und Euler ins Auge gefaßt wurde, vgl. Liouville, *J. éc. polyt.* **21**, **24**, **25** (1832/36), Riemann, *Werke*, S. 332, Lindner, *Berl. Math. Ges.* **7**, 77 (1908) (*Arch. Math. u. Phys.* (3) **13** (1908)).

§ 5. Differentiation von Funktionen mehrerer Veränderlichen, von zusammengesetzten und unentwickelten Funktionen einer Veränderlichen.

Eine Funktion $f(x, y, z, \ldots)$ von mehreren Veränderlichen kann man hinsichtlich jeder einzelnen Variablen differentiieren, indem man die anderen Variablen als konstant ansieht. Man erhält dann die *partiellen Derivierten* (Differentialquotienten) erster Ordnung nach $x, y, z, \ldots$ und bezeichnet sie durch $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}, \ldots$ oder $f_x, f_y, f_z, \ldots$ Weitere Differentiationen liefern die partiellen Derivierten zweiter Ordnung $\frac{\partial^2 f}{\partial x^2}, \frac{\partial^2 f}{\partial x \partial y}, \frac{\partial^2 f}{\partial y \partial x}, \frac{\partial^2 f}{\partial y^2}, \ldots$ oder $f_{xx}, f_{xy}, f_{yx}, f_{yy}, \ldots$ usw. Es besteht der fundamentale Satz von der *Umkehrbarkeit der Differentiationsordnung,* daß nämlich

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$$

ist. Hinreichende Voraussetzungen dabei sind, daß die beiden partiellen Ableitungen 1. Ordnung und eine Ableitung 2. Ordnung vorhanden und stetig sind; es kann auch für *eine* Ableitung 1. Ordnung die Voraussetzung der Stetigkeit entbehrt werden. (H. A. Schwarz, *Vorh. d. Schweiz. naturf. Gesellsch.* 1873, 239 od. *Ges. Abh.* **2**, 275; Timpe, *Math. Ann.* **65**, 310 (1908), woselbst auch weitere Literatur angegeben.)

Infolge dieses Satzes kann man die partiellen Ableitungen n^{ter} Ordnung mit $\frac{\partial^n f}{\partial x^\alpha \partial y^\beta \partial z^\gamma \dots}$ $(\alpha+\beta+\gamma+\dots=n)$ bezeichnen. Eine Funktion von x Veränderlichen hat $\frac{x(x+1)\cdots(x+n-1)}{1\cdot 2\cdots n}$ Ableitungen n^{ter} Ordnung

Die partiellen Derivierten einer *homogenen Funktion*

$$f(x_1, x_2, x_3, \dots)$$

von n Veränderlichen und der Dimension m sind ebenfalls homogen von der Dimension $m-1$, und es ist nach Euler (*Calc. diff.* (1755), 190)

$$x_1 \frac{\partial f}{\partial x_1} + x_2 \frac{\partial f}{\partial x_2} + x_3 \frac{\partial f}{\partial x_3} + \dots + x_n \frac{\partial f}{\partial x_n} = mf.$$

Umgekehrt ist diese Gleichung notwendige und hinreichende Bedingung, daß die Funktion f homogen ist, und m ist dann ihre Dimension. Derselbe Satz liefert für die zweiten Derivierten die Beziehung

$$\sum_{i=1}^{n}\sum_{\varkappa=1}^{n} x_i x_\varkappa \frac{\partial^2 f}{\partial x_i \partial x_\varkappa} = m(m-1)f, \quad \text{usf.}$$

Sind $dx, dy, dz, \dots$ Differentiale der Variablen $x, y, z, \dots$, so sind

$$\partial_x f = f_x dx, \quad \partial_y f = f_y dy, \quad \partial_z f = f_z dz, \quad \dots$$

die *partiellen Differentiale* der Funktion nach den einzelnen Veränderlichen. *Ihre Summe heißt das totale Differential df der Funktion,* also

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy + \frac{\partial f}{\partial z} dz + \cdots$$

Dieser Ausdruck bleibt gültig, auch wenn die Veränderlichen $x, y, z, \dots$ nicht alle voneinander unabhängig sind, d. h. sind $x, y, z, \dots$ ihrerseits Funktionen von $u, v, w, \dots$, so ist

$$\frac{\partial f}{\partial u} du + \frac{\partial f}{\partial v} dv + \frac{\partial f}{\partial w} dw + \cdots = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy + \frac{\partial f}{\partial z} dz + \cdots$$

Sind die ersten partiellen Derivierten stetige Funktionen, so unterscheidet sich das totale Differential um Unendlichkleine höherer Ordnung von dem Zuwachs, den die Funktion erfährt, wenn man die Variablen $x, y, z, \dots$ um $dx, dy, dz, \dots$ ändert.

Betrachtet man $dx, dy, dz, \dots$ als *konstant,* so wird das totale Differential von df als *totales Differential zweiter Ordnung*

von f bezeichnet usf. Das totale Differential n^{ter} Ordnung erhält man durch die symbolische Formel

$$d^n f = [f_x dx]^n.$$

Darin ist $[f_x dx] = f_x dx + f_y dy + f_z dz + \cdots$, und nach erfolgter Potenzierung dieses Ausdrucks ist jedes Produkt

$$f_x^{\alpha} f_y^{\beta} f_z^{\gamma} \ldots \quad (\alpha + \beta + \gamma + \cdots = n)$$

durch die entsprechende Derivierte n^{ter} Ordnung $\dfrac{\partial^n f}{\partial x^{\alpha} \partial y^{\beta} \partial z^{\gamma} \ldots}$ zu ersetzen. Diese Formel setzt aber für $n > 1$ die *Unabhängigkeit* der $x, y, z, \ldots$ voraus.

Sind $x, y, z, \ldots$ Funktionen einer Veränderlichen t, so heißt $f(x, y, z, \ldots)$ eine aus $x, y, z, \ldots$ *zusammengesetzte Funktion* von t. Man kann die Differentialquotienten der Funktion f nach t aus den partiellen Derivierten von f und den Differentialquotienten der $x, y, z, \ldots$ nach t zusammensetzen. Es ist

$$\frac{df}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} + \frac{\partial f}{\partial z}\frac{dz}{dt} + \cdots = \left[\frac{\partial f}{\partial x}\frac{dx}{dt}\right].$$

Die zweite Derivierte ist in symbolischer Schreibweise

$$\frac{d^2 f}{dt^2} = \left[\frac{\partial f}{\partial x}\frac{dx}{dt}\right]^2 + \left[\frac{\partial f}{\partial x}\frac{d^2 x}{dt^2}\right],$$

worin das erste Glied nach dem Potenzieren so, wie eben angegeben, zu behandeln und

$$\left[\frac{\partial f}{\partial x}\frac{d^2 x}{dt^2}\right] = \frac{\partial f}{dx}\frac{d^2 x}{dt^2} + \frac{\partial f}{\partial y}\frac{d^2 y}{dt^2} + \cdots$$

ist.

Sind $x, y, z, \ldots$ unabhängige Funktionen von ebenso vielen anderen Variablen $u, v, w, \ldots$, so ist die Funktionaldeterminante

$$J = \frac{\partial(xyz\ldots)}{\partial(uvw\ldots)}$$

von Null verschieden (vgl. S. 157), und die partiellen Derivierten einer Funktion $f(x, y, z \ldots)$ nach den $x, y, z \ldots$ werden durch die nach $u, v, w \ldots$ in folgender Weise ausgedrückt:

$$\frac{\partial f}{\partial x} = \frac{1}{J}\frac{\partial(fyz\ldots)}{\partial(uvw\ldots)}, \quad \frac{\partial f}{\partial y} = \frac{1}{J}\frac{\partial(xfz\ldots)}{\partial(uvw\ldots)}, \quad \frac{\partial f}{dz} = \frac{1}{J}\frac{\partial(xyf\ldots)}{\partial(uvw\ldots)} \cdots$$

Von besonderer Wichtigkeit für Geometrie und mathematische Physik sind gewisse Verbindungen der partiellen Derivierten

einer Funktion, die gegenüber bestimmten Transformationen *invariant* sind. Es seien von ihnen die *Differentialparameter 1. und 2. Ordnung* einer Funktion von 3 Variablen hervorgehoben:

$$\left(\frac{\partial f}{\partial x}\right)^2+\left(\frac{\partial f}{\partial y}\right)^2+\left(\frac{\partial f}{\partial z}\right)^2 \quad \text{und} \quad \frac{\partial^2 f}{\partial x^2}+\frac{\partial^2 f}{\partial y^2}+\frac{\partial^2 f}{\partial z^2},$$

die bei jeder Koordinatentransformation ungeändert bleiben. Vgl. Lamé, *J. éc. polyt.* **23**, 215 (1834) und seine späteren Lehrbücher z. B. *Leçons s. l. coord. curvilignes* 1859. Beltrami, *Mem. Acc. Bologna* (2) **8** (1869). Die ebenfalls bereits von Lamé gegebene Transformation dieser Differentialausdrücke auf *krummlinige* Koordinatensysteme hat Jacobi (*J. f. Math.* **36**, 117 (1848)) durch Benutzung des Gaußschen Satzes über mehrfache Integrale (vgl. Kap. VIII, § 7) sehr vereinfacht. Vgl. Riemann-Weber, *Part. Differentialgl.*, Braunschweig (1900) **1**, 94. Systematisch tritt der Gesichtspunkt der Invarianz in den Vordergrund in der *absoluten Differentialrechnung* von Ricci und Levi-Civita, *Math. Ann.* **54**, 125 (1901). Vgl. hierzu Kap. XII, § 3.

Durch eine Gleichung $f(x, y) = 0$ ist y *implicit* als Funktion von x gegeben und heißt *unentwickelte Funktion.* Über den Beweis der Existenz der Funktion y und ihrer Derivierten, der für eine ausgedehnte Klasse von Funktionen $f(x, y)$ zuerst von Cauchy gegeben wurde, vgl. etwa Osgood, *Funktionentheorie* Leipzig 1907 (Sammlung Teubner **20**) **1**, 48. Man findet, wenn f_y von Null verschieden ist:

$$\frac{dy}{dx}=-\frac{f_x}{f_y}, \quad \frac{d^2y}{dx^2}=-\frac{f_{xx}f_y^2-2f_{xy}f_xf_y+f_{yy}f_x^2}{f_y^3}.$$

Durch m Gleichungen zwischen $m+1$ Veränderlichen x, y_1, y_2, $\cdots y_m$:

$$f_i(x, y_1, y_2, \cdots y_m) = 0 \qquad i = 1, 2, \ldots m,$$

sind $y_1, y_2, \ldots y_m$ als unentwickelte Funktionen von x definiert, *sobald die Funktionaldeterminante* $J=\frac{\partial(f_1 f_2 \cdots f_m)}{\partial(y_1 y_2 \cdots y_m)}$ *nicht verschwindet* (vgl. S. 157). Ihre Differentialquotienten nach x sind dann

$$\frac{dy_1}{dx}=-\frac{1}{J}\frac{\partial(f_1 f_2 \cdots f_m)}{\partial(x y_2 \ldots y_m)}, \quad \frac{dy_2}{dx}=-\frac{1}{J}\frac{\partial(f_1 f_2 \cdots f_m)}{\partial(y_1 x \ldots y_m)}, \quad \ldots.$$

§ 6. Rollesches Theorem. Mittelwertsätze. Taylorscher Lehrsatz.

Im folgenden wird eine Funktion für solche Intervalle (a, b) der Variablen betrachtet, innerhalb deren *mit Einschluß der Grenzen* die Funktion *stetig* bleibt.

Wenn eine Funktion einer Variablen an den Grenzen eines Intervalls gleiche Werte annimmt und innerhalb des Intervalls überall eine Derivierte besitzt, so verschwindet die Derivierte mindestens in einem Punkt innerhalb des Intervalls. Theorem von Rolle (für ganze rationale Funktionen). Vgl. S. 338.

Es folgt daraus unmittelbar der *Mittelwertsatz*: (Lagrange, *Théor. d. fonct. anal.* 1797; 2. Aufl. 1813 § 39) *Besitzt $f(x)$ für alle Werte eines Intervalls (a, b) eine Derivierte, so gibt es mindestens einen Wert ξ der Variablen innerhalb des Intervalls, so daß*

$$\frac{f(b)-f(a)}{b-a} = f'(\xi)$$

ist. Geometrisch besagt dieser Satz, daß es in jedem endlichen Stück einer stetigen Kurve zu der die Endpunkte verbindenden Sehne mindestens eine parallele Tangente gibt.

Mit $b - a = h$ erhält der Satz die Form: *Es gibt eine Zahl ϑ zwischen* 0 *und* 1, *so daß*

$$f(a+h) - f(a) = hf'(a+\vartheta h). \quad 0 < \vartheta < 1.$$

Allgemeiner, aber ebenfalls eine direkte Folge des Rolleschen Theorems ist der Satz (Genocchi-Peano, *Differentialr.* (1899), 317; der entsprechende Satz für $n+1$ Funktionen und ihre Derivierten bis zur $n-1^{\text{ten}}$ Ordnung ebenda S. 324):

Sind die Funktionen $f(x)$, $\varphi(x)$, $\psi(x)$ in jedem Punkte eines Intervalls differentiierbar, so gibt es mindestens einen Wert ξ innerhalb des Intervalls, so daß

$$\begin{vmatrix} f'(\xi) & \varphi'(\xi) & \psi'(\xi) \\ f(a) & \varphi(a) & \psi(a) \\ f(b) & \varphi(b) & \psi(b) \end{vmatrix} = 0$$

ist. Für $\psi(\xi) = 1$ folgt daraus der Satz von Cauchy (*Calc. diff.* 1829, S. 37):

$$\frac{f(b)-f(a)}{\varphi(b)-\varphi(a)} = \frac{f'(\xi)}{\varphi'(\xi)} \text{ oder } \frac{f(a+h)-f(a)}{\varphi(a+h)-\varphi(a)} = \frac{f'(a+\vartheta h)}{\varphi'(a+\vartheta h)}, \quad 0<\vartheta<1,$$

und dies liefert für $\varphi(x) = x$ den Mittelwertsatz von Lagrange.

Für eine *Funktion von mehreren Veränderlichen* lautet der Mittelwertsatz: *Besitzt eine Funktion* $f(x_1, x_2, x_3, \ldots)$, *deren Variablen der Reihe nach den Intervallen von* a_1 *bis* $a_1 + h_1$, *von* a_2 *bis* $a_2 + h_2$, *von* a_3 *bis* $a_3 + h_3$ *usw. angehören, für alle Werte der Variablen innerhalb dieser Intervalle sämtliche partiellen Derivierten 1. Ordnung, so gibt es eine Zahl* ϑ *zwischen* 0 *und* 1, *so daß*

$$f(a_1 + h_1,\ a_2 + h_2, \ldots) - f(a_1, a_2, \ldots) = \left(h_1 \frac{\partial f}{\partial x_1} + h_2 \frac{\partial f}{\partial x_2} + \cdots\right)_{x_1 = a_1 + \vartheta h_1,\ x_2 = a_2 + \vartheta h_2, \ldots}$$

Sehr wichtige Folgerungen aus dem Mittelwertsatz sind die Sätze:

Eine Funktion, deren Derivierte innerhalb eines Intervalls (a, b) *einen konstanten Wert* k *besitzt, verhält sich in dem ganzen Intervall wie eine lineare Funktion, d. h. es ist für* $a \leqq x \leqq b$:

$$f(x) = f(a) + k(x - a).$$

Wenn die Derivierte einer Funktion innerhalb eines Intervalls beständig Null ist, so besitzt die Funktion im ganzen Intervall einen konstanten Wert.

Über den entsprechenden Satz für Derivierte 2. Ordnung vgl. H. A. Schwarz, *J. f. Math.* **72**, 141 (1870) und den Anhang von Harnack zum 2. Bd. von Serrets Diff.- u. Int.-Rech.

Der *Taylorsche Satz* (Taylor, *Methodus incrementorum* 1715) ist zunächst eine Verallgemeinerung des Mittelwertsatzes: *Wenn eine Funktion* $f(x)$ *in einem Intervall von* a *bis* $a + h$ *mit Einschluß der Grenzen endlich ist und alle Differentialquotienten bis zur* n^{ten} *Ordnung besitzt, so ist*

$$f(a + h) = f(a) + \frac{h}{1!} f'(a) + \frac{h^2}{2!} f''(a) + \cdots + \frac{h^{n-1}}{(n-1)!} f^{(n-1)}(a) + R_n,$$

wobei der Rest R_n *mit Hilfe einer im ganzen Intervall endlichen und differentiierbaren Funktion* $\varphi(x)$, *deren Derivierte im Intervall nicht verschwindet, durch die Gleichung*

$$R_n = \frac{\varphi(a + h) - \varphi(a)}{\varphi'(a + \vartheta h)} \frac{(1 - \vartheta)^{n-1} h^{n-1}}{1 \cdot 2 \cdot 3 \cdots (n-1)} f^{(n)}(a + \vartheta h) \quad 0 < \vartheta < 1$$

bestimmt ist.

Diese sehr allgemeine Form des Restes stammt von Schlömilch (*Handb. d. Diff.- u. Integralr.* 1847).

Für $\varphi(x) = (a + h - x)^p$ mit einem beliebigen *positiven* Exponenten p ergibt sich

$$R_n = \frac{h^n(1-\vartheta)^{n-p}}{p(n-1)!} f^{(n)}(a + \vartheta h),$$

und wenn man hier $p = n$ oder $p = 1$ nimmt, erhält man $R_n = \frac{h^n}{n!} f^{(n)}(a+\vartheta h)$, Lagrange, *Théorie d. fonct. anal.* (1797) § 40; $R_n = \frac{h^n(1-\vartheta)^{n-1}}{(n-1)!} f^{(n)}(a + \vartheta h)$, Cauchy, *Résumé* (1823), 173.

Am kürzesten wird der Taylorsche Satz mit Hilfe der *partiellen Integration* abgeleitet (Prony 1805 nach Angabe von Cauchy, *Résumé* 1823); dabei erhält man den *Rest* in *exakter* Form, d. h. ohne die unbestimmte Größe ϑ, als *bestimmtes Integral* (Lagrange, *Théor. d. fonct. anal.*):

$$R_n = \frac{1}{(n-1)!} \int_0^h f^{(n)}(x + t)(h - t)^{n-1} dt,$$

woraus sich die oben angeführten Ausdrücke für R_n ableiten lassen.

Der Taylorsche Satz für $a = 0$, $h = x$ heißt *Maclaurinscher Lehrsatz.* Maclaurin, *Treatise of fluxions* 1742, jedoch bereits von Taylor gegeben.

Besitzt die im Intervall a bis $a + h$ stetige Funktion $f(x)$

1. *innerhalb des Intervalls nur endliche Derivierte von beliebig hoher Ordnung*, und ist

2. $\lim_{n=\infty} R_n = 0$ für jedes zwischen 0 und 1 liegende ϑ,

so besteht für $f(a + h)$ die nach Potenzen von h fortschreitende *Taylorsche Entwicklung*

$$f(a + h) = f(a) + hf'(a) + \frac{h^2}{1 \cdot 2} f''(a) + \cdots.$$

Diese *beiden* Bedingungen sind für das Bestehen der Entwicklung *hinreichend*, nicht aber — wie Lagrange (*Th. d. fonct.* chap. V, § 30) vermutete — die erste allein. Ebensowenig ist dessen Ansicht richtig, daß die Reihe, sobald sie konvergiere, den Wert $f(a + h)$ ergebe; bereits Cauchy (*Calc. inf.* 1823, 152 u. 299) hat darauf hingewiesen, daß z. B. in

der (reellen) Umgebung von $x = 0$ die Funktionen $f(x)$ und $f(x) + e^{-\frac{1}{x^2}}$ die gleiche Taylorentwicklung besitzen. Überzeugendere Beispiele bei du Bois-Reymond, *Math. Ann.* **21**, 114 (1876) und Pringsheim, *Münch. Ber.* (1892), 211 oder *Math. Ann.* **42**, 153 (1893). Die abschließende Arbeit von Pringsheim, *Math. Ann.* **44**, 57 (1894) gibt die *notwendigen* und *hinreichenden* Bedingungen für die Darstellbarkeit einer Funktion durch die Taylorsche Reihe, wobei die erste Bedingung im wesentlichen bestehen bleibt, während die *zweite* Bedingung durch die weniger umfassende ersetzt werden kann, *daß mit wachsendem n das Cauchysche Restglied für alle in Betracht kommenden Werte von ϑ gleichmäßig gegen Null konvergiert.* Vgl. hierzu Pascal, *Esercizi e note critiche di calc. infinit.* (1895), 176 ff., W. H. Young, *Quart. Journ.* **40**, 157 (1909).

Eine erschöpfende Behandlung der Taylorschen Reihe erfordert die Heranziehung *komplexer* Werte der Variablen und ist Sache der *Funktionentheorie* (s. d.)

Eine eingehende Darstellung der Geschichte des Taylorschen Satzes hat Pringsheim, *Bibl. Math.* (3) **1**, 433 (1900) gegeben.

Für *Funktionen mehrerer Veränderlichen* lautet der *Taylorsche Satz: Besitzt eine Funktion $f(x_1, x_2, x_3 \ldots)$, deren Variable der Reihe nach den Intervallen $(a_1, a_1 + h_1)$, $(a_2, a_2 + h_2)$, $(a_3, a_3 + h_3)$ usw. angehören, für alle Werte der Variablen innerhalb dieser Intervalle sämtliche Derivierten bis zur n^{ten} Ordnung, so ist in symbolischer Schreibweise*

$$f(a_1 + h_1, a_2 + h_2, \cdots) = f(a_1, a_2, \cdots)$$
$$+ \left[h \frac{\partial f}{\partial a}\right] + \frac{1}{2!}\left[h \frac{\partial f}{\partial a}\right]^2 + \frac{1}{3!}\left[h \frac{\partial f}{\partial a}\right]^3 + \cdots + \frac{1}{(n-1)!}\left[h \frac{\partial f}{\partial a}\right]^{n-1} + R_n.$$

Darin ist $\left[h \frac{\partial f}{\partial a}\right] = h_1 \frac{\partial f}{\partial a_1} + h_2 \frac{\partial f}{\partial a_2} + \cdots$, und die Potenzen dieses Ausdruckes sind in derselben Weise zu behandeln wie im vorigen Paragraphen; das Restglied R_n kann man in verschiedener Form geben, am einfachsten:

$$R_n = \frac{1}{n!}\left[h \frac{\partial f}{\partial x}\right]^n_{\substack{x_1 = a_1 + \vartheta h_1 \\ x_2 = a_2 + \vartheta h_2 \\ \cdots\cdots\cdots}} \qquad (0 < \vartheta < 1).$$

§ 7. Grenzwerte von Ausdrücken in unbestimmter Form.

Eine Funktion von der Form $F(x) = \frac{f(x)}{\varphi(x)}$ erscheint *unbestimmt*, wenn für $x = a$ entweder beide Funktionen $f(x)$ und $\varphi(x)$ *Null* oder beide *unendlich* werden. Unter dem Wert der Funktion $F(x)$ an dieser Stelle versteht man den *Grenzwert*, wenn ein solcher existiert, den der Quotient für $x = a$ besitzt. Die Funktionen $f(x)$ und $\varphi(x)$ mögen in der *Umgebung* von a stetig und differentiierbar sein. Dann gelten die Sätze:

1. *Wenn $f(x)$ und $\varphi(x)$ für $x = a$ stetig sind und dort verschwinden, so ist*

$$\lim_{x=a} \frac{f(x)}{\varphi(x)} = \lim_{x=a} \frac{f'(x)}{\varphi'(x)}.$$

2. *Wenn $f(x)$ und $\varphi(x)$ für $x = a$ unendlich werden,* wenn ferner $\varphi'(x)$ an dieser Stelle von Null verschieden ist und in der Umgebung das Vorzeichen nicht ändert, so ist ebenfalls

$$\lim_{x=a} \frac{f(x)}{\varphi(x)} = \lim_{x=a} \frac{f'(x)}{\varphi'(x)}.$$

Wenn man bei Anwendung dieser Sätze wieder zu Ausdrücken der Form $\frac{0}{0}$ oder $\frac{\infty}{\infty}$ gelangt, so geht man zu höheren Differentialquotienten weiter; es kann aber vorkommen, daß man mit dem Verfahren nicht zu Ende kommt und gleichwohl ein Grenzwert für den Quotienten existiert. (Vgl. Pascal, *Esercizi e note crit.* 240.)

Der erste Satz geht auf Joh. Bernoulli (1694, Brief an l'Hospital), der zweite auf Euler (*Inst. calc. diff.* (1755), 738) zurück. Die strenge Behandlung beginnt mit Cauchy; abschließend bei Stolz, *Math. Ann.* **14**, 231 (1879).

Andere unbestimmte Formen, wie $0 \cdot \infty$, $\infty - \infty$, 0^0, ∞^0, 1^∞ können stets auf die oben behandelten zurückgeführt werden. Die folgenden Beispiele zeigen Eigenschaften der darin vorkommenden Funktionen, von denen oft Gebrauch zu machen ist.

$$\lim_{x=0} \frac{\sin x}{x} = 1, \quad \lim_{x=0} \frac{\operatorname{tg} x}{x} = 1,$$

$$\lim_{x=\infty} x^\alpha e^{-mx} = 0 \quad \text{und} \quad \lim_{x=\infty} \frac{(\ln x)^m}{x^\alpha} = 0$$

für jedes positive α und m.

Wenn sich die Funktionen $f(x)$ und $\varphi(x)$ in der Umgebung von a, d. h. für $x = a + h$ in *Reihen* nach Potenzen von h entwickeln lassen, so ist es oft vorteilhafter, an Stelle des obigen Verfahrens diese Reihenentwicklungen einzusetzen, mit der niedrigsten im Nenner vorkommenden Potenz von h zu kürzen und zur Grenze $h = 0$ überzugehen.

Über unbestimmte Ausdrücke bei Funktionen von *mehreren* Veränderlichen vgl. Genocchi-Peano, *Differentialr.* (1899), 170 ff.

§ 8. Wachsende und abnehmende Funktionen. Maxima und Minima.

Eine Funktion $f(x)$ heißt im Punkte x *wachsend*, wenn sich ein Intervall, zu dem x gehört, angeben läßt, so daß innerhalb des Intervalls mit jedem $x_1 > x$ auch $f(x_1) > f(x)$, dagegen mit jedem $x_0 < x$ auch $f(x_0) < f(x)$ ist; die Funktion heißt *abnehmend* in x, wenn innerhalb eines ebensolchen Intervalls mit jedem $x_1 > x$ zugleich $f(x_1) < f(x)$, dagegen mit jedem $x_0 < x$ zugleich $f(x_0) > f(x)$ ist.

Ist $f'(x)$ positiv, so ist $f(x)$ wachsend, ist $f'(x)$ negativ, so ist $f(x)$ abnehmend.

Wenn für den Wert x alle Derivierten bis zur $(n-1)^{\text{ten}}$ Ordnung verschwinden und die erste nicht verschwindende Derivierte $f^{(n)}(x)$ *ungerader* Ordnung ist, so ist $f(x)$ wachsend oder abnehmend, je nachdem $f^{(n)}(x)$ positiv oder negativ ist. Ist aber $f^{(n)}(x)$ von *gerader* Ordnung, so nimmt die Funktion im Punkt x weder zu noch ab, was geometrisch durch eine zur x-Achse parallele Tangente gekennzeichnet ist.

Eine Funktion $f(x)$ hat im Punkte x ein *Maximum*, wenn die Funktionswerte an den zu x benachbarten Stellen sämtlich *kleiner* als $f(x)$ sind, d. h. wenn sich ein Intervall, *innerhalb* dessen x liegt, angeben läßt, so daß für irgendeinen nicht mit x zusammenfallenden Punkt x_1 des Intervalls $f(x_1) < f(x)$ ist. Dagegen hat $f(x)$ in x ein *Minimum*, wenn die Funktionswerte an den zu x benachbarten Stellen sämtlich *größer* als $f(x)$ sind, d. h. wenn für jedes nicht mit x identische x_1 des Intervalls $f(x_1) > f(x)$ ist. Als gemeinsamer Name für Maximum und Minimum hat sich die Bezeichnung *Extremum* eingebürgert.

Soll $f(x)$ im Punkt x ein Extremum haben, so muß dort die erste Derivierte Null und die erste nicht verschwindende

Derivierte $f^n(x)$ von gerader Ordnung sein, und zwar wird $f(x)$ ein Maximum oder Minimum, je nachdem $f^{(n)}(x)$ einen negativen oder positiven Wert hat.

In der Sache übereinstimmend, bei Anwendungen manchmal brauchbarer ist folgendes Kriterium: *Bei einem Extremum von $f(x)$ wechselt $f'(x)$ das Vorzeichen; geht es dabei (bei wachsendem x) von positiven zu negativen Werten, so hat $f(x)$ ein Maximum, andernfalls ein Minimum.*

Über die Fälle, in denen dieser Satz versagt, z. B. $f(x)$ nicht differentiierbar, mit einseitigen Differentialquotienten u. a. vgl. Pascal, *Esercizi etc.* 215 ff., Stolz, *Grundz. d. Diff.- u. Int.-Rech.* (1893) **1**, 206.

Eine Funktion von *mehreren Veränderlichen* hat an einer Stelle $(x_1, x_2, \ldots)$ ein *Maximum* oder *Minimum*, wenn der Wert der Funktion an dieser Stelle *größer* oder *kleiner* ist als sämtliche Funktionswerte in der Umgebung; sind also α_i, β_i $(i = 1, 2, 3 \ldots)$ positive Größen und $h_1, h_2, \ldots$ irgendwelche Zahlen, die den Ungleichungen $-\alpha_i < h_i > \beta_i$ genügen und nicht alle gleichzeitig Null sind, so muß bei einem

$$\textit{Maximum:} \quad f(x_1 + h_1, x_2 + h_2, \ldots) - f(x_1, x_2, \ldots) < 0,$$

$$\textit{Minimum:} \quad f(x_1 + h_1, x_2 + h_2, \ldots) - f(x_1, x_2, \ldots) > 0$$

sein.

Für ein Extremum der Funktion $f(x_1, x_2, \ldots x_n)$ ist *notwendige* Bedingung, *daß die ersten partiellen Differentialquotienten sämtlich verschwinden* und daß die niedrigste Ordnung, bei der nicht alle partiellen Differentialquotienten Null sind, eine *gerade* Zahl ist.

Sei k diese Zahl, so hängt die Entscheidung, ob ein *Maximum* oder *Minimum* stattfindet, von dem Verhalten des im Taylorschen Satz auftretenden symbolischen Ausdrucks

$$\left[h \frac{\partial f}{\partial x}\right]^k = F$$

ab. Dieser stellt eine Form k^{ten} Grades in $h_1, h_2, \ldots, h_n$ vor. Für solche Formen (mit geradem k) gelten die Begriffe *definit*, *indefinit* und *semidefinit*, wie sie auf S. 123 für *quadratische* Formen erläutert sind. Ein Extremum ist sicher vorhanden, wenn die Form *definit* ist, und zwar ein *Maximum*, wenn sie beständig *negativ*, ein *Minimum*, wenn sie beständig *positiv* ist. Ist F *indefinit*, so ist die betreffende Stelle kein Extremum der

Funktion. Ist aber die Form *semidefinit*, d. h. verschwindet sie für Werte $x_1, x_2, \ldots, x_n$, die nicht alle Null sind, während sie für alle übrigen Wertesysteme dasselbe Vorzeichen hat, so erfordert die Entscheidung, ob ein Extremum stattfindet, spezielle Untersuchungen, wegen derer auf Genocchi-Peano, *Differentialr.*, Scheeffer, *Math. Ann.* **35**, 541 (1885), Stolz, *Grundzüge* **1**, 211 ff. verwiesen sei.

Sei nun $k = 2$, also mögen für eine Stelle $(x_1, x_2, \ldots, x_n)$ alle ersten, aber nicht alle zweiten Differentialquotienten Null sein, so ist F die quadratische Form

$$F = \sum_{i=1}^{n} \sum_{\varkappa=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_k} h_i h_k = \sum \sum f_{ik} h_i h_k,$$

wenn $\frac{\partial^2 f}{\partial x_i \partial x_k} = f_{ik}$ gesetzt wird. Bildet man die Reihe der Determinanten

$$D_\alpha = \begin{vmatrix} f_{11} & f_{12} & \cdots & f_{1\alpha} \\ f_{21} & f_{22} & \cdots & f_{2\alpha} \\ \vdots & & & \\ f_{\alpha 1} & f_{\alpha 2} & \cdots & f_{\alpha\alpha} \end{vmatrix} \quad \text{für } \alpha = 1, 2, \ldots, n,$$

so ist F *positiv definit*, wenn *alle* D_α *positiv*, dagegen *negativ definit*, wenn die D_α *abwechselnd negativ oder positiv* (mit *negativem* $D_1 = f_{11}$) sind.

Bei einer Funktion von *zwei* Variablen ist F *definit*, wenn $D_2 = \frac{\partial^2 f}{\partial x_1^2} \frac{\partial^2 f}{\partial x_2^2} - \left(\frac{\partial^2 f}{\partial x_1 \partial x_2}\right)^2$ *positiv* ist und die Funktion besitzt alsdann ein *Maximum* oder *Minimum*, je nachdem $\frac{\partial^2 f}{\partial x_1^2}$ $\left(\text{und damit auch } \frac{\partial^2 f}{\partial x_2^2}\right)$ *negativ* oder *positiv* ist. (Lagrange, *Misc. Taur.* **1** (1759).) Ist $D_2 < 0$, so ist F *indefinit* und kein Extremum vorhanden; der *semidefinite* Fall ist durch $D_2 = 0$ gekennzeichnet.

Unter einem *bedingten Maximum* oder *Minimum (relative Extreme)* versteht man ein Extremum der Funktion $f(x_1, x_2, \ldots x_n)$, wobei zwischen den Variablen $k < n$ Bedingungsgleichungen $\varphi_1 = 0$, $\varphi_2 = 0$, $\ldots$, $\varphi_k = 0$ zu erfüllen sind. Man führt mit Lagrange (*Th. d. fonct.* 268) zweckmäßig *Multiplikatoren* λ_1, $\lambda_2, \ldots, \lambda_k$ ein und bildet die n Gleichungen

$$\frac{\partial f}{\partial x_i} + \sum_{\alpha=1}^{k} \lambda_\alpha \frac{\partial \varphi_\alpha}{\partial x_i} = 0. \qquad (i = 1, 2, \ldots n)$$

Diese müssen im Fall eines Extremums erfüllt sein und liefern zusammen mit den gegebenen k Gleichungen Werte für $x_1 \ldots x_n$, $\lambda_1 \ldots \lambda_k$. Ob aber diesen Lösungen wirklich Extreme entsprechen und welcher Art sie sind, ist durch weitere Untersuchung zu entscheiden. (Vgl. Stolz, *Grundzüge* **1**, 245.) Diese Methode ist in der *Ausgleichungsrechnung* bei der Ausgleichung bedingter Beobachtungen von Wichtigkeit. (Korrelatenmethode. Gauß, *Theor. comb. observ. Suppl.* § 11 (1826).)

Kapitel VIII.

Integralrechnung.

Von *Paul Epstein* in Straßburg i. E.

§ 1. Das unbestimmte Integral. Definition und allgemeine Sätze.

Die Integralrechnung geht davon aus, daß eine gegebene Funktion aufgefaßt wird als *Differentialquotient einer zu bestimmenden Funktion.* Ist $f(x)$ gegeben, so ist eine Funktion $F(x)$ so zu bestimmen, daß $F'(x) = f(x)$ oder $f(x)\,dx$ *das Differential von* $F(x)$ ist. Man nennt dann $F(x)$ *das Integral von* $f(x)$ oder auch nach Lagrange (*Théor. d. fonct. anal.*) *primitive Funktion von* $f(x)$ und schreibt

$$F(x) = \int f(x)\,dx.$$

Die Funktion $f(x)$ heißt *Integrand*; dieser ist also der Differentialquotient des Integrals. Besitzt $f(x)$ ein Integral $F(x)$, so besitzt es unendlich viele, die man durch Hinzufügen von *willkürlichen Konstanten* zu $F(x)$ erhält. Man kann daher dem Integral, um es völlig festzulegen, in einem Punkt a einen bestimmten Wert A vorschreiben; es ist dann

$$\int f(x)\,dx = F(x) - F(a) + A.$$

Eine in einem Intervall stetige Funktion besitzt dort ein Integral, welches für einen Punkt a des Intervalls einen willkürlich vorgeschriebenen Wert A annimmt.

Weitergehende Bedingungen für die Existenz eines Integrals kommen in der Lehre von den bestimmten Integralen zur Sprache.

Es bestehen die allgemeinen Sätze:

1. *Einen konstanten Faktor des Integranden kann man vor das Integralzeichen nehmen.*

2. *Das Integral aus einer Summe von Funktionen ist gleich der Summe der Integrale aus den einzelnen Funktionen.*

3. Eine *unendliche Reihe* von stetigen Funktionen läßt sich in einem Intervall, in dem sie *gleichmäßig konvergiert*, *gliedweise integrieren*, und die Reihe der Integrale konvergiert in demselben

Intervall ebenfalls gleichmäßig. Aber auch bei nicht gleichmäßiger Konvergenz kann gliedweise Integration zulässig sein. Vgl. Osgood, *Am. Journ.* **19**, 155 (1894), Arzelà, *Rend. Acc. Linc.* (4) **1**, 321 (1885), Schoenflies, *Bericht üb. die Mengenlehre*, *Math.-Ver.* **8** (1900) Kap. VII, Borel, *Leç. s. l. fonct. de var. réelles* (1905), 45, Vitali, *Rend. circ. Palermo* **23**, 137 (1907).

Insbesondere läßt sich eine *Potenzreihe*, sobald sie konvergiert, gliedweise integrieren.

4. Sei $f(u, v)$ eine aus zwei Funktionen u, v von x zusammengesetzte Funktion und es sei das Integral $\int f(u, v)\,dx$ auszuführen. Läßt sich nun die Integration ausführen, *wenn man v als konstant ansieht*, so möge dieses *„partielle Integral"* mit

$$\int_u f(u, v)\,dx = y$$

bezeichnet werden. Dann ist

$$\int f(u, v)\,dx = \int_u f(u, v)\,dx - \int \frac{\partial y}{\partial v}\,dv$$

$$= \int_u f(u, v)\,dx - \int \frac{\partial y}{\partial v}\frac{dv}{dx}\,dx.$$

Vgl. hierzu Worpitzky, *Lehrb. d. Diff.- u. Int.-Rech.*, Berlin 1880, 73, Brendel, *Math. Ann.* **55**, 248 u. 599 (1902), Hatzidakis, ebenda **57**, 134 (1903).

Gewöhnlich versteht man unter *partieller Integration* einen speziellen Fall dieses Satzes, wenn nämlich $f(u, v) = u \cdot v$ ist. Dann ist

$$\int uv\,dx = v\int u\,dx - \int\left(\frac{dv}{dx}\int u\,dx\right)dx,$$

oder wenn $u = \dfrac{dw}{dx} = w'$ ist:

$$\int vw'\,dx = vw - \int wv'\,dx.$$

Dies Verfahren ist anzuwenden, wenn man einen Faktor des Integranden als Differentialquotienten einer Funktion w erkennt.

5. Durch *Einführung einer neuen Variablen (Substitution)* vermöge der Gleichung $x = \psi(z)$ verwandelt sich das Integral $\int f(x)\,dx$ in $\int f[\psi(z)]\psi'(z)\,dz$.

Ein Integral heißt *ausführbar*, wenn man es durch einen geschlossenen Ausdruck mit einer endlichen Anzahl von *elementaren* Funktionen darstellen kann. Man versucht es dabei mit Hilfe von partieller Integration und Substitution neuer Variabler auf eins der folgenden *Grundintegrale* (bei denen die willkürliche Konstante weggelassen ist) zurückzuführen:

$$\int x^n dx = \frac{x^{n+1}}{n+1} \ (n \neq -1), \quad \int \frac{dx}{x} = \log x, \quad \int e^x dx = e^x,$$

$$\int \sin x \, dx = -\cos x, \quad \int \cos x \, dx = \sin x,$$

$$\int \frac{dx}{\sqrt{1-x^2}} = \arcsin x, \quad \int \frac{dx}{1+x^2} = \operatorname{arc\,tg} x.$$

§ 2. Integration wichtiger Klassen von Funktionen.

I. *Integration rationaler Funktionen.* Um eine rationale Funktion $\frac{P(x)}{Q(x)}$, bei der man $P(x)$ von niedrigerem Grad als $Q(x)$ voraussetzen darf, zu integrieren, zerlegt man sie in *Partialbrüche* (s. Algebra S. 262). Es treten dann für jede Wurzel a der Gleichung $Q(x) = 0$ Integrale von folgendem Typus auf:

$$\int \frac{dx}{x-a} = \log (x-a)$$

und

$$\int \frac{dx}{(x-a)^\varkappa} = -\frac{1}{\varkappa - 1} \cdot \frac{1}{(x-a)^{\varkappa-1}} \quad (\varkappa > 1).$$

Unter diesen Wurzeln können auch komplexe sein; wenn aber $Q(x)$ nur reelle Koeffizienten hat, sind die komplexen Wurzeln *paarweise konjugiert,* und die zwei derartigen Wurzeln $\alpha + i\beta$ und $\alpha - i\beta$ entsprechenden Integrale lassen sich zu solchen vom Typus

$$\int \frac{A(x-\alpha)+B}{(x-\alpha)^2+\beta^2} dx = A \log \sqrt{(x-\alpha)^2+\beta^2} + \frac{B}{\beta} \operatorname{arc\,tg} \frac{x-\alpha}{\beta}$$

und

$$\int \frac{Ax+B}{(x^2+2px+q)^\varkappa} dx \quad (\varkappa > 1)$$

zusammenfassen, wobei $(x-\alpha)^2 + \beta^2 = x^2 + 2px + q$ ist. Die letzteren Integrale werden durch partielle Integration auf solche von niedrigerem Exponenten $\varkappa$ zurückgeführt. Setzt man nämlich $\int \frac{dx}{(x^2+2px+q)^\varkappa} = J_\varkappa$ und $p^2 - q = \Delta$, so ist

$$\int \frac{Ax+B}{(x^2+2px+q)^x}\,dx = -\frac{A}{2(x-1)(x^2+2px+q)^{x-1}} + (B-Ap)J_x$$

und

$$J_x = -\frac{1}{2\Delta(x-1)}\left\{\frac{x+p}{(x^2+2px+q)^{x-1}} + (2x-3)J_{x-1}\right\}$$

Man hat demnach den Satz: *Das Integral einer rationalen Funktion besteht im allgemeinen aus einem rationalen und einem transzendenten Teil. Der rationale Bestandteil rührt von den mehrfachen Wurzeln von $Q(x)=0$ her. In dem transzendenten Teil treten folgende Funktionen auf:*

1) *für jede reelle Wurzel a von $Q(x)=0$ die Funktion $\log(x-a)$,*

2) *für jedes Paar konjugiert komplexer Wurzeln $\alpha \pm i\beta$ die Funktionen $\log((x-\alpha)^2+\beta^2)$ und $arc\,tg\,\frac{x-\alpha}{\beta}$.*

Die Berechnung des rationalen Teils kann auf *rationalem Wege* ohne Kenntnis der Wurzeln erfolgen, wenn man die Zerlegung

$$\frac{P(x)}{Q(x)} = \frac{d}{dx}\left(\frac{P_1(x)}{Q_1(x)}\right) + \sum_h \frac{C_h}{x-a_h} + \sum_i \frac{A_i(x-\alpha_i)+B_i}{(x-\alpha_i)+\beta_i^2}$$

zugrunde legt. Darin ist

$$Q_1(x) = (x-\xi_1)^{e_1-1}(x-\xi_2)^{e_2-1}\ldots(x-\xi_m)^{e_m-1},$$

wenn $\xi_1 \ldots \xi_m$ die *mehrfachen* Wurzeln von $Q(x)=0$ und $e_1 \ldots e_m$ ihre Multiplizitäten bedeuten, und daher auch $Q_1(x)$ *der größte gemeinsame Teiler von $Q(x)$ und seiner Derivierten.* (Vgl. Baltzer, *Leipz. Ber.* (1873), 535, Hermite, *Cours d'anal.* (1873), 265, Stolz, *Grundzüge* **1**, 285.)

II. *Integrale algebraischer Funktionen.* Durch eine Gleichung

$$f_0(x)y^m + f_1(x)y^{m-1} + \cdots + f_m(x) = 0,$$

deren Koeffizienten *ganze rationale* Funktionen von x sind, wird y als *algebraische Funktion* von x definiert. Jede in x und y *rationale* Funktion $R(x,y)$ ist ebenfalls algebraische Funktion von x und die Gesamtheit dieser Funktionen bildet eine *Klasse* oder einen *Körper* algebraischer Funktionen. (Weber, *Algebra* **1**, 491.) Die Integrale solcher Funktionen $\int R(x,y)\,dx$ heißen *Abelsche Integrale.* Zur Klassifikation der algebraischen Funktionen und ihrer Integrale dient der Begriff des *Geschlechts.* Es wird am anschaulichsten geometrisch definiert. Jeder Funktion der Klasse entspricht eine *algebraische Kurve* und ist n ihre

Ordnung, d die Anzahl ihrer *Doppel-* und *Rückkehrpunkte*, so ist das *Geschlecht* $p = \frac{1}{2}(n-1)(n-2) - d$ eine für alle Kurven der Klasse konstante Zahl. Ist das Geschlecht *Null*, so lassen sich die Koordinaten eines jeden Kurvenpunktes als *rationale* Funktionen eines Parameters t darstellen (rationalisieren), und die Kurve ist eine *rationale* oder *unikursale* Kurve. (Clebsch, *J. f. Math.* **64**, 43 (1864), Cayley, *Lond. Math. Soc.* 1865 Okt.) Folglich kann *jedes Abelsche Integral vom Geschlecht* $p = 0$ *auf das Integral einer rationalen Funktion zurückgeführt und mithin durch elementare Funktionen ausgedrückt werden.* Dagegen sind die Integrale von algebraischen Funktionen, deren Geschlecht $p > 0$ ist, nicht ausführbar und führen zu neuen Transzendenten; ihre Theorie wird eingehend in den Kapiteln 18, 19 und 20 behandelt.

Die einfachsten algebraischen Funktionen vom *Geschlecht Null* werden durch eine *quadratische Gleichung* erzeugt, *deren Diskriminante eine lineare oder quadratische Funktion von* x *ist.* Man kann ohne Beschränkung der Allgemeinheit eine *reine* quadratische Gleichung, also die Funktion $y = \sqrt{ax^2 + 2bx + c}$ zugrunde legen.[1]) Die hieraus entspringende Klasse algebraischer Funktionen $R(x, y)$ kann auf unendlich viele Arten rationalisiert werden. Ist nämlich $ax^2 + 2bx + c = f(x)$ und *sei für irgendein* $x = \xi$ *der Wert von* $y = \eta$, *also* $\eta^2 = f(\xi)$, *so ist*

$$(1) \qquad x = \xi + \frac{f'(\xi) + 2t\eta}{t^2 - a}, \quad y = -\eta + t\frac{f'(\xi) + 2t\eta}{t^2 - a}$$

eine rationale Parameterdarstellung von x *und* y. Es wird also durch eine Substitution (1) jedes Integral der Klasse $\int R(x, y)\,dx$ in das Integral einer *rationalen* Funktion übergeführt. Dabei wird $dx = -\frac{2y}{t^2 - a}\,dt$, und nach Ausführung der Integration ist $t = \frac{y+\eta}{x-\xi} = \frac{a(x+\xi)+2b}{y-\eta}$ zu setzen. Sind alle im Integranden vorkommenden Größen reell, so läßt sich ξ immer so wählen, daß auch die Substitution (1) reell ist.

Sind α, β die Nullpunkte der Funktion $f(x)$, also $f(x) = a(x-\alpha)(x-\beta)$ und nimmt man $\xi = \alpha$, so liefert (1) die einfachste Parameterdarstellung

---

1) Ist $y = \sqrt{ax+b}$ $\left(\text{oder allgemein } y = \sqrt[m]{\frac{ax+b}{cx+\partial}}\right)$, so nimmt man in den Integralen $\int R(x, y)\,dx$ einfach y als Integrationsvariable.

$$x = \frac{\alpha t^2 - \beta}{t^2 - 1}, \quad y = t\,\frac{\alpha - \beta}{t^2 - 1},$$

und es ist

$$t = \sqrt{\frac{x - \beta}{x - \alpha}}.$$

Als wichtige Beispiele dieser Integrale seien die folgenden angeführt, wobei $f(x) = ax^2 + 2bx + c$, $\Delta = b^2 - ac$, $\varphi(x, \xi) = ax\xi + b(x + \xi) + c$ ist:

$$\int \frac{dx}{\sqrt{f(x)}} = \frac{1}{\sqrt{a}} \log\left(ax + b + \sqrt{af(x)}\right), \qquad a > 0$$

$$= -\frac{1}{\sqrt{-a}} \arcsin \frac{ax + b}{\sqrt{\Delta}}, \qquad a < 0$$

$$\int \sqrt{f(x)}\,dx = \frac{ax + b}{2a}\sqrt{f(x)} - \frac{\Delta}{2a\sqrt{a}} \log\left(ax + b + \sqrt{af(x)}\right), \quad a > 0$$

$$= \frac{ax + b}{2a}\sqrt{f(x)} + \frac{\Delta}{2a\sqrt{-a}} \arcsin \frac{ax + b}{\sqrt{\Delta}}, \qquad a < 0$$

$$\int \frac{dx}{(x - \xi)\sqrt{f(x)}} = -\frac{1}{\sqrt{f(\xi)}} \log \frac{\varphi(x, \xi) + \sqrt{f(x)\,f(\xi)}}{x - \xi}, \qquad f(\xi) > 0$$

$$= \frac{1}{\sqrt{-f(\xi)}} \arcsin \frac{\varphi(x, \xi)}{(x - \xi)\sqrt{\Delta}}, \qquad f(\xi) < 0.$$

Dagegen

$$\int \frac{dx}{(x - \xi)\sqrt{f(x)}} = -\frac{\sqrt{f(x)}}{(x - \xi)\sqrt{\Delta}},$$

wenn $f(\xi) = 0$ ist.

Auf das erste und die beiden letzten Integrale lassen sich alle Integrale der Klasse durch Partialbruchzerlegung zurückführen. Vgl. die auf Weierstraß zurückgehende Darstellung bei Stolz, *Grundzüge* **1**, 312; **2**, 126, sowie die für die allgemeine Theorie der Abelschen Integrale grundlegende Abhandlung von Aronhold, *J. f. Math.* **61**, 95 (1863).

Das zuletzt angeführte Integral ist eine *algebraische* Funktion und zwar von derselben Klasse wie der Integrand. Das entspricht einem allgemeinen Satz von Abel, *J. f. Math.* **4**, 264. Vgl. Laplace, *Théor. anal. d. probabilités* 3. éd. (1820) S. 7, Liouville, *J. éc. polyt.* **14**, 149.

Ein Integral von der Form $\int x^p (ax^q + b)^r dx = J(p, q, r)$, in dem p, q, r *rationale* Zahlen sind, heißt ein *binomisches Integral*. Es können dabei p und q als *ganze* Zahlen vorausgesetzt werden, was nötigenfalls durch eine Substitution $x = y^n$ erreicht wird, und überdies darf q als *positiv* angenommen werden.

Ein binomisches Integral ist in den folgenden Fällen ausführbar: 1. *wenn* r, 2. *wenn* $\frac{p+1}{q}$, 3. *wenn* $\frac{p+1}{q}+r$ *eine ganze Zahl ist.*

In den beiden letzten Fällen wird das Integral durch die Substitution $ax^q+b=y$ bezw. $a+bx^{-q}=y$ in das Integral einer *rationalen* Funktion übergeführt. (Newton 1676. *Opuscula* **1**, 335.) Tschebyscheff (*J. d. math.* **18**, 87 (1853)) hat bewiesen, daß diese Fälle die einzigen sind, in denen sich das binomische Integral auf algebraische und logarithmische Funktionen zurückführen läßt.

Jedes binomische Integral kann auf ein solches reduziert werden, bei dem $0 \leqq r < 1$ ist und p *und* q *ganze Zahlen* sind, so daß $0 \leqq p < q$. Dazu dienen die durch partielle Integration gewonnenen Formeln:

$$\left(\frac{p+1}{q}+r\right)J(p,q,r)=\frac{x^{p+1}}{q}(ax^q+b)^r+brJ(p,q,r-1),$$

$$\left(\frac{p+1}{q}+r\right)aJ(p,q,r)=\frac{x^{p-q+1}}{q}(ax^q+b)^{r+1}$$
$$-\left(\frac{p+1}{q}-1\right)bJ(p-q,q,r).$$

III. *Die Integration transzendenter Funktionen* sucht man durch geeignete Substitutionen auf diejenige von *rationalen* Funktionen zurückzuführen. Das gelingt in folgenden Fällen, wobei R das Zeichen für eine *rationale* Funktion bedeutet.

1. $\int R(e^{ax})dx$ durch die Substitution $e^{ax}=y$, also $dx=\frac{dy}{ay}$.

Hierher gehören Integrale, wie:

$$\int e^{ax}\sin bx\,dx=e^{ax}\frac{a\sin bx-b\cos bx}{a^2+b^2},$$

$$\int e^{ax}\cos bx\,dx=e^{ax}\frac{a\cos bx+b\sin bx}{a^2+b^2}.$$

2. $\int R(\sin x,\cos x)dx$ durch die Substitution $\operatorname{tg}\frac{x}{2}=y$, also

$$\sin x=\frac{2y}{1+y^2},\quad \cos x=\frac{1-y^2}{1+y^2},\quad dx=\frac{2dy}{1+y^2}.$$

So ist z. B.

$$\int \frac{dx}{\sin x} = \log \operatorname{tg} \frac{x}{2}, \quad \int \frac{dx}{\cos x} = \log \operatorname{tg} \left(\frac{\pi}{4} + \frac{x}{2}\right),$$

$$\int \frac{dx}{a \sin x + b \cos x} = \frac{1}{\sqrt{a^2 + b^2}} \log \operatorname{tg} \left(\frac{\pi}{4} + \frac{x - \lambda}{2}\right),$$

$$\text{worin } \lambda = \operatorname{arc\,tg} \frac{b}{a},$$

$$\int \frac{dx}{a \cos^2 x + b \sin^2 x} = \frac{1}{\sqrt{ab}} \operatorname{arc\,tg} \left(\sqrt{\frac{b}{a}} \cdot \operatorname{tg} x\right),$$

$$\int \frac{dx}{a \cos^2 x - b \sin^2 x} = \frac{1}{2\sqrt{ab}} \log \frac{\sqrt{a} + \sqrt{b} \operatorname{tg} x}{\sqrt{a} - \sqrt{b} \operatorname{tg} x}.$$

Auf die letzten Integrale werden

$$\int \frac{dx}{a + b \cos x} = 2 \int \frac{d\left(\frac{x}{2}\right)}{(a + b) \cos^2 \frac{x}{2} + (a - b) \sin^2 \frac{x}{2}}$$

und

$$\int \frac{dx}{a + b \cos x + c \sin x} = \int \frac{d(x - \lambda)}{a + r \cos (x - \lambda)} \quad \left(r = \sqrt{b^2 + c^2},\ \lambda = \operatorname{arc\,tg} \frac{c}{b}\right)$$

zurückgeführt.

Die Integrale von der Form $\int \sin^m x \cos^n x\, dx$ werden durch Substitutionen, wie $\sin x = y$ oder $\operatorname{tg} x = z$, in *binomische* Integrale verwandelt. Auch kann man, wenn $m + n$ nicht Null ist, durch die Formeln

$$\int \sin^m x \cos^n x\, dx = \frac{\sin^{m+1} x \cos^{n-1} x}{m + n} + \frac{n - 1}{m + n} \int \sin^m x \cos^{n-2} x\, dx$$

$$= -\frac{\sin^{m-1} x \cos^{n+1} x}{m + n} + \frac{m - 1}{m + n} \int \sin^{m-2} x \cos^n x\, dx$$

die Exponenten im Integranden auf Werte zwischen -1 und $+1$ reduzieren.

3. Bedeutet P ein *Polynom* und sind $\alpha, \beta, \ldots \mu$ irgendwelche Zahlen, so ist das Integral $\int P(x, e^{\alpha x}, e^{\beta x}, \ldots e^{\mu x})\, dx$ *stets ausführbar,* denn es besteht aus einer endlichen Anzahl von Integralen der Form

$$\int x^n e^{ax}\,dx = \frac{d^n}{da^n}\left(\frac{e^{ax}}{a}\right)$$

$$= \frac{(-1)^{n-1}n!}{a^{n+1}}\,e^{ax}\left[1 - ax + \frac{(ax)^2}{2!} - \frac{(ax)^3}{3!} + \cdots \pm \frac{(ax)^n}{n!}\right].$$

Ein Integral $\int e^x R(x)\,dx$ wird durch Partialbrüche in eine Anzahl Integrale von der Form $A^m \int \frac{e^x\,dx}{(x-a)^{m+1}}$ zerlegt, und diese werden vermöge der Formel

$$\int \frac{e^x\,dx}{(x-a)^{m+1}} = -\frac{e^x}{m(x-a)^m} + \frac{1}{m}\int \frac{e^x\,dx}{(x-a)^m}$$

schließlich auf *eine mit e^x multiplizierte rationale Funktion* und ein *Aggregat von Integralen der Form* $A\int \frac{e^x\,dx}{x-a}$ reduziert. Die letzteren lassen sich nicht durch elementare Funktionen ausdrücken; sie werden aber sämtlich auf eine *einzige neue transzendente* Funktion $\int \frac{e^z}{z}\,dz$, den *Integrallogarithmus* $li(e^z)$ zurückgeführt. Ebenso führen die Integrale von der Form $\int \sin x\,R(x)\,dx$ und $\int \cos x R(x)\,dx$ zu den Transzendenten $\int \frac{\sin z}{z}dx$ und $\int \frac{\cos z}{z}\,dz$.

Vgl. zu diesem Paragraphen: Hermite, *Cours d'analyse* 1873, S. 261—356, Hardy, *The integration of functions of a single variable*, *Cambridge tracts in Mathematics* 1905.

§ 3. Das bestimmte Integral. Definitionen.

I. Der Übergang vom unbestimmten zum bestimmten Integral wird am einfachsten durch die geometrische Anschauung vermittelt. Nimmt man auf der Kurve $y = f(x)$ einen festen Punkt mit der Abszisse $x = a$, so ist der *Inhalt der Fläche* von der zugehörigen Ordinate bis zu der Ordinate irgendeines Kurvenpunktes mit der Abszisse x eine gewisse Funktion von x, und man zeigt, daß der *Differentialquotient dieser Funktion gleich der Endordinate*, *also gleich* $f(x)$ ist. Daraus folgt, daß der gesamte Flächeninhalt durch das Integral $\int f(x)\,dx$ angegeben wird, aber darin ist die Konstante so zu bestimmen, daß das Integral für $x = a$ den Wert Null hat. Man bezeichnet

alsdann das Integral durch $\int_a^x f(x)\,dx$, und wenn $F(x)$ einen Wert des *unbestimmten* Integrals bedeutet $\left(\text{also } \frac{dF}{dx} = f(x)\right)$, so ist *das bestimmte Integral zwischen den Grenzen a und x* definiert durch

$$(1) \qquad \int_a^x f(x)\,dx = F(x) - F(a).$$

II. Zu einer anderen Definition des bestimmten Integrals gelangt man, wenn man den Begriff des *Flächeninhalts*, der zunächst nur für geradlinig begrenzte Flächenstücke Bedeutung hat, genauer zu präzisieren sucht. Hierauf beruhen die Definitionen von Cauchy (*J. éc. pol.* **19**, 571 (1823)) und die umfassendere von Riemann (*Hab.-Schr.* Gött. 1854 = *Werke*, 239), nämlich:

Es sei $f(x)$ in dem endlichen Intervall (a, b) eine *begrenzte* Funktion[1]). Man zerlegt das Intervall in irgendeiner Weise in *Teilintervalle* $d_1, d_2, \ldots d_n$ und es sei x_i der Wert von x an einer *beliebigen* Stelle des Teilintervalls d_i. Man bildet dann die Summe

$$J = d_1 f(x_1) + d_2 f(x_2) + \cdots + d_n f(x_n)$$

und betrachtet ihr Verhalten, während man die Anzahl der Teilintervalle beständig so vermehrt, daß alle gegen Null konvergieren. *Wenn dann diese Summe einen bestimmten Grenzwert besitzt, der unabhängig ist von der Art der Teilung des Intervalls* (a, b) *und von der Wahl der Punkte* x_i, *so nennt man* $f(x)$ *integrierbar, und der Grenzwert heißt das bestimmte Integral von* $f(x)$ *zwischen den Grenzen a und b.* Man schreibt

$$(2) \qquad \int_a^b f(x)dx = \lim\left[d_1 f(x_1) + d_2 f(x_2) + \cdots + d_n f(x_n)\right].$$

Diese Auffassung des Integrals als *Grenzwert einer Summe* ist älter als die auf der Umkehrung der Differentialrechnung beruhende; sie geht allgemein auf Leibniz (*Mscr.* vom 29. X. 1675) zurück, findet sich aber in einzelnen Beispielen — entsprechend dem aufs Konkrete gerichteten Charakter der griechischen Mathematik, die den Begriff der Funktion nicht kannte —

1) *Fonction bornée*, d. h. die Funktion bleibt immer zwischen zwei endlichen Zahlen A und B.

bereits bei Archimedes (287—212 v. Chr.) in der Schrift über die *Quadratur der Parabel* und namentlich in der erst 1906 gefundenen *Abhandlung von den mechanischen Lehrsätzen* (vgl. Heiberg und Zeuthen, *Bibl. Math.* (3) **7**, 321 (1907))

Die folgenden Sätze lassen die Tragweite und den Geltungsbereich der Riemannschen Definition erkennen:

1. Sei D_i die *Schwankung* der Funktion $f(x)$ innerhalb des Teilintervalls d_i (vgl. S. 34), so ist *notwendige und hinreichende Bedingung*, *damit* $f(x)$ *integrabel ist*, daß die Summe $\sum d_i D_i$ den *Grenzwert Null* besitzt.

2. Damit $f(x)$ *integrierbar* ist, ist *notwendig* und *hinreichend*, daß die *Summe derjenigen Teilintervalle*, in denen die Schwankung von $f(x)$ größer als eine positive Zahl ε ist, *bei wachsendem n beliebig klein wird.*

3. Damit $f(x)$ *integrierbar* ist, ist *notwendig* und *hinreichend*, daß die *Menge* ihrer *Unstetigkeitspunkte* innerhalb (a, b) den *Inhalt Null* (vgl. S. 29) besitzt.

Insbesondere sind folgende Funktionen integrierbar:

a) Alle innerhalb (a, b) *endlichen* und *stetigen* Funktionen.

b) Alle innerhalb (a, b) endlichen und *abteilungsweise monotonen* Funktionen (vgl. Differentialrechnung § 2).

c) Die innerhalb (a, b) endlichen Funktionen mit einer *endlichen* Anzahl von *Unstetigkeiten* und diejenigen *punktiert unstetigen* Funktionen (vgl. S. 37), deren Unstetigkeitspunkte eine *endliche Anzahl von Häufungsstellen* besitzen (allgemeiner eine *abzählbare Menge* bilden).

d) Diejenigen innerhalb (a, b) endlichen und *punktiert unstetigen* Funktionen, bei denen die *Anzahl der Sprünge*, die größer sind als eine beliebig kleine positive Zahl ε, stets *endlich* ist und nur bei unbeschränkter Abnahme von ε über jedes Maß hinaus wachsen kann.

Der Satz von Hankel (*Math. Ann.* **20**, 92 (1870)), daß *jede* innerhalb (a, b) *punktiert unstetige* Funktion dort integrierbar sei, ist *nicht* richtig (Smith, *Lond. Math. Soc.* **6**, 148 (1875), Volterra, *Giorn. di Math.* **19**, 76 (1881)), während umgekehrt jede integrierbare Funktion innerhalb des Integrationsintervalls nur punktiert unstetig werden kann.

4. Eine integrierbare Funktion darf man innerhalb des Integrationsintervalls in den Punkten einer Punktmenge erster Gattung[1]) *beliebig abändern, ohne daß sich der Wert des Integrals ändert.*

1) d. h. eine Punktmenge mit einer *endlichen* Anzahl von Ableitungen (vgl. S. 28).

5. *Stetige Funktionen von integrierbaren Funktionen sind wieder integrierbar.*

Eine unmittelbare Folge der Definition ist der Satz: *Der Quotient* $\frac{1}{b-a}\int_a^b f(x)\,dx$ *kann als der Grenzwert des arithmetischen Mittels* $\frac{1}{n}\sum_1^n f(x_i)$ *der n Werte der Funktion in beliebigen Teilpunkten des Intervalls* (a, b) *angesehen werden.* Über die hieran anknüpfenden Näherungsmethoden zur Berechnung bestimmter Integrale *(mechanische Quadratur)* vgl. *Differenzenrechnung* § 4.

Von den an Riemann anknüpfenden und seine Ideen weiterführenden Arbeiten seien genannt: Darboux, *Ann. éc. norm.* (2) 4, (1875), Thomae, *Einl. in d. Theor. d. best. I.* (1875), Ascoli, *Acc. Linc. Atti* (1876), 863. Vgl. die Darstellungen in Dini, *Grundlagen* und Jordan, *Cours d'analyse.*

III. Neuerdings ist der Integralbegriff durch Lebesgue bedeutend erweitert worden, und zwar auf Grund des Begriffs des *Inhalts* einer Punktmenge (S. 29). Während bei Riemann ein Intervall (a, b) der *unabhängigen Variablen* in Teilintervalle zerlegt wird, zerlegt Lebesgue die *Gesamtschwankung* der *Funktion* innerhalb (a, b). Sei also $y = f(x)$ eine *begrenzte* Funktion, m ihre *untere*, M ihre *obere* Grenze innerhalb (a, b); wählt man zwischen m und M eine Reihe von wachsenden Funktionswerten, so daß

$$m = y_0 < y_1 < y_2 \cdots < y_{n-1} < y_n = M$$

ist, so bilden die Werte von x, für die $y_{i-1} < f(x) < y_i$ und ebenso diejenigen, für die $y_i = f(x)$ ist, jedesmal eine *Punktmenge.* Wenn nun jede dieser Punktmengen *meßbar* ist, und zwar *bei jeder beliebigen Wahl der Zwischenwerte* $y_1, \ldots y_{n-1}$, so heißt die Funktion $f(x)$ *meßbar innerhalb* (a, b). Sei nun e_i der *Inhalt* der zu $y_{i-1} < f(x) < y_i$ gehörigen, e_i' der *Inhalt* der zu $y_i = f(x)$ gehörigen *Punktmenge*; wenn dann die Summen

$$(3) \qquad S = \sum_1^n e_i y_i + \sum_1^n e_i' y_i \quad \text{und} \quad s = \sum_1^n e_i y_{i-1} + \sum_1^n e_i' y_i$$

bei unbegrenzter Vermehrung von n unabhängig von der Wahl der Zwischenwerte y_i *zu derselben Grenze* J konvergieren, so

heißt $f(x)$ *summierbar*, und J ist das *Integral* von $f(x)$ im Intervall (a, b). Es gilt der Satz:

Damit eine begrenzte Funktion *summierbar* ist, ist *notwendig* und *hinreichend*, daß sie *meßbar ist.*

Der Lebesguesche Integralbegriff ist umfassender als der Riemannsche und enthält ihn als besonderen Fall, denn der Begriff der *meßbaren* Funktionen ist weiter als der der *integrierbaren.* Man kennt *keine begrenzte* Funktion, die *nicht summierbar* ist; insbesondere sind *alle unstetigen Funktionen irgendeiner Baireschen Klasse* (vgl. S. 37) *summierbar.*

Lebesgue hat seinen Integralbegriff auch auf *nicht begrenzte* Funktionen ausgedehnt; dabei zeigt sich aber der merkwürdige Umstand, daß es unter diesen Funktionen solche gibt, die im gewöhnlichen Sinn *integrierbar*, aber *nicht summierbar* sind.

Über die Theorie von Lebesgue vgl. Lebesgue, *Ann. di mat.* (3) **7**, 231 (1902) und dess. Verf. *Leç. s. l'intégration*, Paris 1904; Borel, *Leç. s. l. fonct. de var. réelles*, Paris 1905, Schönflies, *Entwickl. d. Lehre v. d. Punktmannigf.* **2**, 318 (1908).

Stellt man das bestimmte Integral in der Definition von Riemann oder Lebesgue an die Spitze der Integralrechnung, so wird das *unbestimmte Integral* als $\int_a^x f(x)\,dx$ mit einer willkürlichen unteren Grenze a definiert, d. h. es wird das bestimmte Integral als *Funktion der oberen Grenze* betrachtet. Es bestehen die folgenden Sätze, durch die der Zusammenhang zwischen der *ersten* und der Riemannschen Definition hergestellt wird:

1. *Die Integralfunktion* $\int_a^x f(x)\,dx$ *ist eine endliche und stetige Funktion von* x.

2. *Ist auch* $f(x)$ *eine stetige Funktion, so besitzt die Integralfunktion in jedem Punkt* x *eine bestimmte Derivierte, deren Wert gleich* $f(x)$ *ist.*

3. *Fundamentalsatz der Integralrechnung. Falls die stetige Funktion* $F(x)$ *in jedem Punkt eine bestimmte endliche Derivierte* $f(x)$ *besitzt, so ist*

$$\int_a^x f(x)\,dx = F(x) - F(a).$$

Es ist aber die erste Definition des bestimmten Integrals *nicht völlig äquivalent* mit der Riemannschen, denn einerseits gibt es Funktionen, die im Riemannschen Sinn *integrabel* sind, aber keine *primitive* Funktionen besitzen, andererseits gibt es *primitive* Funktionen, deren Derivierte *nicht* im Riemannschen Sinn integrabel sind. Vgl. Du Bois-Reymond, *Math. Ann.* **16**, 115 (1880), Pasch, ebenda **30**, 153 (1887), Volterra, *Giorn. di Mat.* **19**, 76 (1881), Scheeffer, *Acta Mat.* **5**, 183 u. 279 (1884), Hahn, *Monatsh. f. Math.* **16**. 161 (1905).

Uneigentliche bestimmte Integrale. Die Riemannsche Definition versagt in folgenden Fällen:

1. wenn $f(x)$ im Integrationsintervall *Unendlichkeitsstellen* besitzt,

2. wenn das Integrationsintervall *unendlich groß* ist.

Wenn es nun gelingt, durch geeignete Festsetzungen das Integral $\int_a^b f(x)\,dx$ auch in diesen Fällen zu definieren, so nennt man ein solches Integral ein *uneigentliches Integral.* Cauchy, *J. éc. polyt.* **19**, 572 (1823).

Im folgenden sei $a < b$ vorausgesetzt, was nach dem ersten Satz im folgenden Paragraphen stets erreichbar ist.

Es möge also die Funktion $f(x)$ in einem Punkt c des Intervalls (a, b) *unendlich* werden, aber in den Teilintervallen $(a, c-\delta)$ und $(c+\varepsilon, b)$ integrierbar sein, *wie klein man auch die positiven Zahlen δ und ε annehmen mag;* man definiert dann

$$\int_a^b f(x)\,dx = \lim_{\delta=0} \int_a^{c-\delta} f(x)\,dx + \lim_{\varepsilon=0} \int_{c+\varepsilon}^b f(x)\,dx,$$

vorausgesetzt, daß diese Grenzwerte existieren, endlich und *unabhängig sind von der Art, wie δ und ε gegen Null konvergieren.* Wenn $f(x)$ in einer endlichen Anzahl von Punkten $c_1 c_2 \ldots c_n$ unendlich, im übrigen aber integrierbar ist, so wird ebenso

$$\int_a^b f(x)\,dx = \lim_{\delta_1=0} \int_a^{c_1-\delta_1} + \lim_{\substack{\varepsilon_1=0\\ \delta_2=0}} \int_{c_1+\varepsilon_1}^{c_2-\delta_2} + \lim_{\substack{\varepsilon_2=0\\ \delta_3=0}} \int_{c_2+\varepsilon_2}^{c_3-\delta_3} + \cdots + \lim_{\varepsilon_n=0} \int_{c_n+\varepsilon_n}^b$$

definiert. Wird $f(x)$ in unendlich vielen Punkten von (a, b) mit einer *endlichen Anzahl von Häufungsstellen* unendlich und schneidet man die Umgebungen $\omega_1, \omega_2, \ldots \omega_\varkappa$ aller Häufungsstellen heraus, so zerfällt (a, b) in $\varkappa + 1$ Teilintervalle, deren

jedes nur eine endliche Anzahl von Unendlichkeitsstellen enthält. Wenn nun auf Grund der letzten Definition das Integral über jedes Teilintervall existiert, wie klein man auch $\omega_1, \omega_2, \ldots \omega_\varkappa$ wählen mag, so wird die Summe dieser Integrale für $\lim \omega_1 = 0, \ldots \lim \omega_\varkappa = 0$ als $\int_a^b f(x)\,dx$ definiert. Entsprechend kann man zu Definitionen des Integrals kommen, wenn $f(x)$ in den Punkten einer *Punktmenge erster Gattung* unendlich wird. (Vgl. Dini, *Grundlagen* S. 406.)

Es kann vorkommen, daß im allgemeinen keine bestimmten Grenzwerte existieren und dennoch sich für $\int_a^b f(x)\,dx$ ein bestimmter Wert ergibt, *wenn man eine Beziehung zwischen den gegen Null konvergierenden Größen δ und ε annimmt;* (Beispiel $\int_{-1}^{+1} \frac{dx}{x} = \ln a$, wenn $\delta = a \cdot \varepsilon$). Für den Fall eines einzigen Unendlichkeitspunktes c im Intervall (a, b) nennt Cauchy den aus der Annahme $\delta = \varepsilon$ sich ergebenden Grenzwert, falls er existiert, den *Hauptwert* des Integrals. Auch Dirichlet hat hiervon Gebrauch gemacht (*Vorl. üb. bestimmte I.* herausgeg. von Arendt 1904), jedoch ist die Notwendigkeit der Einführung dieses Begriffes bestritten worden. (Riemann, *Werke,* S. 225, Kronecker, *Vorl. üb. die Theor. d. Integrale* S. 210.)

Wird das *Integrationsintervall unendlich,* indem entweder die *obere* oder die *untere Grenze* oder *beide unendlich groß* werden, so definiert man

$$\int_a^\infty f(x)\,dx = \lim_{b=\infty} \int_a^b f(x)\,dx,$$

$$\int_{-\infty}^b f(x)\,dx = \lim_{a=-\infty} \int_a^b f(x)\,dx,$$

$$\int_{-\infty}^\infty f(x)\,dx = \lim_{\substack{a=-\infty \\ b=+\infty}} \int_a^b f(x)\,dx.$$

Dabei muß im letzten Fall der Grenzwert unabhängig sein von der *Reihenfolge* und der *Art des Unendlichwerdens* von a und b. Auch hier kann es vorkommen, daß er nur unter Annahme

einer bestimmten Beziehung zwischen a und b existiert; insbesondere ist der Grenzwert $\lim_{a=\infty} \int_{-a}^{+a} f(x)\,dx$, falls er vorhanden ist, der *Hauptwert* des Integrals.

Wenn die angegebenen Grenzwerte für ein uneigentliches Integral existieren und endlich sind, so heißt das Integral *konvergent*; es heißt *absolut konvergent*, wenn auch das mit dem *absoluten Wert der Funktion* gebildete Integral konvergiert, *bedingt konvergent*, wenn dies nicht der Fall ist.

Es bestehen die Sätze: 1. Ist $f(x)$ im Endpunkt b des endlichen Intervalls (a, b) unendlich, sonst aber im ganzen Intervall integrierbar, so ist zur absoluten Konvergenz von $\int_a^b f(x)\,dx$ *hinreichend* die Existenz einer positiven Zahl $\mu < 1$, so daß $\lim_{x=b-0} (x-b)^\mu f(x)$ existiert und endlich ist.

2. Wenn unter den gleichen Voraussetzungen überdies $f(x)$ in (a, b) sein Zeichen nicht wechselt, so ist $\int_a^b f(x)\,dx$ sicher *divergent, sobald* $|(x-b)f(x)|$ *im ganzen Intervall und auch für* $x = b$ *größer als eine bestimmte positive Zahl bleibt.*

3. Zur absoluten Konvergenz von $\int_a^\infty f(x)\,dx$ ist *hinreichend* die Existenz einer Zahl $\mu > 1$, so daß $\lim_{x=\infty} x^\mu f(x)$ existiert und endlich ist.

4. Wenn $f(x)$ im Intervall (a, ∞) sein Zeichen nicht wechselt, so ist $\int_a^\infty f(x)\,dx$ sicher *divergent, sobald* $|x\,f(x)|$ *im ganzen Intervall und auch für* $x = \infty$ *größer als eine bestimmte positive Zahl bleibt.*

Vgl. hierzu Riemann, *Werke*, 229, Du Bois-Reymond, *J. f. Math.* **76**, 88 (1872), Pringsheim, *Math. Ann.* **37**, 591 (1890), Dini, *Grundlagen*, Kap. 17 u. 18, Stolz, *Grundzüge* **1**, 401. Weitere Literatur bei Pascal, *Esercizi e note critiche* 272.

Durch den folgenden Satz wird für eine wichtige Klasse von Integralen die Konvergenz (nicht aber die *absolute* Konvergenz) gesichert:

Ist $\varphi(x)$ eine Funktion von der Art, daß das Integral $\int_a^x \varphi(x)\,dx$ mit unendlich wachsendem x in endlichen Grenzen bleibt (z. B. $\varphi(x) = \sin \alpha x$ oder $\cos \alpha x$), so ist zur Konvergenz des Integrals $\int_a^\infty f(x)\,\varphi(x)\,dx$ hinreichend, daß die Funktion $f(x)$ im Integrationsintervall endlich bleibt und daß ihr absoluter Wert bei wachsendem x beständig zu Null abnimmt.

Schließlich sei der Satz erwähnt: Ist das Integral $\int_0^a f(x)\,dx$ konvergent und ist

1. $f(x)$ *gerade* Funktion, so ist $\int_{-a}^{+a} f(x)\,dx = 2\int_0^a f(x)\,dx$,

2. $f(x)$ *ungerade* Funktion, so ist $\int_{-a}^{+a} f(x)\,dx = 0$.

§ 4. Haupteigenschaften der bestimmten Integrale.

Die allgemeinen Sätze über unbestimmte Integrale (§ 1) sind sinngemäß auf die bestimmten Integrale zu übertragen. Dazu kommen noch folgende Sätze:

1. *Bei Vertauschung der Grenzen ändert ein bestimmtes Integral nur das Vorzeichen:*

$$\int_a^b f(x)\,dx = -\int_b^a f(x)\,dx.$$

2. *Gehören a, b, c oder allgemeiner $a, b, c_1, c_2, \ldots c_n$ einem Intervall an, in dem $f(x)$ integrierbar ist, so ist*

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx$$

und

$$\int_a^b f(x)\,dx = \int_a^{c_1} f(x)\,dx + \int_{c_1}^{c_2} f(x)\,dx + \cdots \int_{c_{n-1}}^{c_n} f(x)\,dx + \int_{c_n}^b f(x)\,dx.$$

3. Ist m die *untere*, M die *obere* Grenze von $f(x)$ im Intervall (a, b), so ist stets

$$m \leqq \frac{1}{b-a} \int_a^b f(x)\,dx \leqq M.$$

4. *Erster Mittelwertsatz.* Sind $f(x)$ und $\varphi(x)$ im Intervall (a, b) integrierbar *und behält dort* $\varphi(x)$ *sein Zeichen*, so ist

$$m \int_a^b \varphi(x)\,dx \lessgtr \int_a^b f(x)\varphi(x)\,dx \lessgtr M \int_a^b \varphi(x)\,dx.$$

Ist $f(x)$ *stetige* Funktion, so gibt es einen Funktionswert $f(\xi)$ in einem bestimmten Punkt ξ des Intervalls (a, b), so daß

$$\int_a^b f(x)\varphi(x)\,dx = f(\xi) \int_a^b \varphi(x)\,dx.$$

Cauchy-Moigno, *Calc. intégr.* **2**, 48 (1844), Dirichlet, *Werke* **1**, 137. Übertragungen auf *komplexes Gebiet* von Darboux, *J. de math.* (2) **3**, 294 (1876) und Weierstraß (vgl. Hermite, *Cours d'analyse* 4. éd. 1891, S. 62).

5. *Zweiter Mittelwertsatz.* Sind $f(x)$ und $\varphi(x)$ im Intervall (a, b) integrierbar *und ist* $f(x)$ *dort monoton*, so ist

$$\int_a^b f(x)\,\varphi(x)\,dx = f(a) \int_a^\xi \varphi(x)\,dx + f(b) \int_\xi^b \varphi(x)\,dx.$$

Bonnet, *J. de math.* **14**, 249 (1849), Du Bois-Reymond, *J. f. Math.* **69**, 81 (1869), Hölder, *Gött. Anz.* (1894), 519, Pringsheim, *Münch. Ber.* (1900), 209, G. Kowalewski, *Math. Ann.* **60**, 151 (1905). Eine eingehende Untersuchung über diesen Satz mit Heranziehung *komplexer Größe* enthält Brunn, *Beziehungen des Du Bois-Reymondschen Mittelwertsatzes zur Ovaltheorie* 1905.

Einen Mittelwertsatz, der — allerdings unter spezielleren Voraussetzungen für die Funktionen $f(x)$ und $\varphi(x)$ — engere Grenzen liefert als der erste, hat Fejér, *Math. Ann.* **64**, 287 (1907) gegeben.

6. Ist $f(x)$ eine Funktion, für die $\int_a^b f(x)\,dx$ absolut konvergiert, $\varphi(x, s)$ eine innerhalb (a, b) stetige Funktionen von x,

die dort überall unter einer endlichen positiven Zahl bleibt und für $s = \alpha$ gleichmäßig gegen einen endlichen Grenzwert $\varphi(x)$ konvergiert, so konvergieren auch die Integrale $\int_a^b f(x)\varphi(x,s)\,dx$ und $\int_a^b f(x)\varphi(x)\,dx$ absolut, und es ist

$$\lim_{s=\alpha}\int_a^b f(x)\varphi(x,s)\,dx = \int_a^b f(x)\varphi(x)\,dx.$$

7. Eine Funktion $f(x,s)$ von zwei Veränderlichen sei in einem Bereiche $(a \leqq x \leqq b,\ p \leqq s \leqq q)$ als Funktion von x *integrierbar*, als Funktion von s *stetig; alsdann stellt das bestimmte Integral*

$$\int_a^b f(x,s)\,dx$$

eine im Intervall (p,q) — welches auch unendlich groß sein kann — *stetige Funktion von* s *dar.* Diese Funktion kann in einem zum Bereich gehörigen Intervall (α,β) von s *unter dem Integralzeichen integriert* und, sobald auch $\frac{\partial f(x,s)}{\partial s}$ in dem angegebenen Bereiche *stetig* ist, *unter dem Integralzeichen nach* s *differentiiert werden,* d. h. es ist

$$\int_\alpha^\beta\Big[\int_a^b f(x,s)\,dx\Big]ds = \int_a^b\Big[\int_\alpha^\beta f(x,s)\,ds\Big]dx,$$

$$\frac{d}{ds}\int_a^b f(x,s)\,dx = \int_a^b \frac{\partial f(x,s)}{\partial s}\,dx.$$

Sind aber die Grenzen a und b ebenfalls von s abhängig, so tritt an Stelle des letzten Satzes die Gleichung

$$\frac{d}{ds}\int_a^b f(x,s)\,dx = \int_a^b \frac{\partial f(x,s)}{\partial s}\,dx + f(b,s)\frac{db}{ds} - f(a,s)\frac{da}{ds}.$$

Zu diesen Sätzen vgl. Stolz, *Grundzüge* **1**, 425, Pascal, *Esercizi e note critiche* 299, Arzelà, *Acc. Linc. Rend.* (4) **1** (1885) u. (5) **6**, 290 (1897), Osgood, *Ann. of Math.* (2) **3**, 129 (1902) u. (2) **9**, 119 (1908).

Sämtliche Sätze dieses Paragraphen lassen sich auf *uneigentliche Integrale* übertragen, wenn man nur beachtet, daß alle vorkommenden Integrale *konvergent* sein müssen.

§ 5. Zusammenstellung einiger bestimmter Integrale.

Eine auch nur einigermaßen vollständige Übersicht der wichtigeren bestimmten Integrale kann hier nicht gegeben werden. Es sei außer auf die bekannten Lehrbücher der Differential- und Integralrechnung besonders auf die Vorlesungen von Dirichlet (herausgeg. von G. F. Meyer, Leipzig 1871 und Arndt, Braunschweig 1904), Kronecker (herausgeg. von Netto, Leipzig 1894) und Thomae (Leipzig 1908) verwiesen. Die vollständigste Darstellung der Methoden zur Ermittlung bestimmter Integrale findet man bei Bierens de Haan, *Exposé de la théorie ... des intégr. déf.*, Amsterdam 1862, und die reichhaltigste Sammlung von solchen Integralen in desselben Verfassers *Tables d'intégr. déf.*, ebd. 1858, 2. éd. 1867. Vgl. dazu Lindman, *Examen des nouv. tables ... de B. de H* Stockholm 1891.

I. Als die einfachsten bestimmten Integrale sind diejenigen anzusehen, bei denen die *unbestimmte* Integration ausführbar ist; man findet dann den Wert des bestimmten Integrals auf Grund der ersten Definition § 3. So ist

$$(1)\quad \int_0^1 x^a\,dx = \frac{1}{a+1},\ a>-1. \qquad \int_1^\infty x^a\,dx = -\frac{1}{a+1},\ a<-1.$$

$$(2)\quad \int_0^\infty e^{-ax}\,dx = \frac{1}{a},\quad a>0.$$

Durch n-malige *Differentiation* dieses Integrals nach a folgt

$$\int_0^\infty e^{-ax}x^n\,dx = (-1)^n\int_0^1 x^{a-1}(\log x)^n\,dx = \frac{n!}{a^{n+1}}$$

und durch *Integration* desselben Integrals nach a im Intervall $(1, a)$:

$$\int_0^\infty \frac{e^{-x}-e^{-ax}}{x}\,dx = \log a.$$

$$(3)\quad \int_0^1 \frac{dx}{\sqrt{1-x^2}} = \frac{\pi}{2}, \qquad \int_0^{\frac{1}{2}} \frac{dx}{\sqrt{1-x^2}} = \frac{\pi}{6}.$$

(4) $$\int_0^\infty \frac{dx}{1+x^2} = \frac{\pi}{2}; \quad \int_0^1 \frac{dx}{1+x^2} = \int_1^\infty \frac{dx}{1+x^2} = \frac{\pi}{4}.$$

(5) $$\int_0^{\frac{\pi}{2}} \frac{dx}{a\cos^2 x + b\sin^2 x} = \frac{\pi}{2\sqrt{ab}}, \qquad a, b \text{ positiv.}$$

Durch Differentiation dieses Integrals nach a und b:

$$\int_0^{\frac{\pi}{2}} \frac{dx}{(a\cos^2 x + b\sin^2 x)^2} = \frac{\pi}{4\sqrt{ab}}\left(\frac{1}{a} + \frac{1}{b}\right).$$

(6) $$\int_0^{\frac{\pi}{2}} \frac{dx}{a + b\cos x} = \frac{1}{\sqrt{a^2 - b^2}} \arccos \frac{b}{a}, \qquad a^2 > b^2$$

$$= \frac{1}{\sqrt{b^2 - a^2}} \log \frac{b + \sqrt{b^2 - a^2}}{a}, \qquad a^2 < b^2$$

$$= \frac{1}{a}, \qquad a = b.$$

(7) $$\int_0^\pi \frac{\cos nx\, dx}{1 - 2a\cos x + a^2} = \frac{\pi a^n}{1 - a^2}, \quad |a| < 1,$$

$$= \frac{\pi a^{-n}}{a^2 - 1}, \quad |a| > 1.$$

n ganze positive Zahl,

(8) $$\int_0^\infty e^{-ax} \cos bx\, dx = \frac{a}{a^2 + b^2},$$

$$\int_0^\infty e^{-ax} \sin bx\, dx = \frac{b}{a^2 + b^2}.$$

$a > 0$.

Durch Integration des ersten Integrals nach b von 0 bis b (Euler 1781):

$$\int_0^\infty e^{-ax} \frac{\sin bx}{x}\, dx = \operatorname{arc\,tg} \frac{b}{a}.$$

Hieraus

$$\int_0^\infty \frac{\sin bx}{x}\, dx = \pm \frac{\pi}{2}, \quad \text{je nachdem } b \gtrless 0$$

und

$$\int_0^\infty \frac{\cos ax \sin x}{x} dx = \frac{\pi}{2}, \quad \text{wenn } |a|<1,$$

$$= \frac{\pi}{4}, \quad \text{wenn } |a|=1,$$

$$= 0, \quad \text{wenn } |a|>1.$$

Dirichlets *diskontinuierlicher Faktor* (*Berl. Ber.* 1839, 18 u. *Abhandl.* 1841, 61). Bereits 1811 neben anderen diskontinuierlichen Funktionen von Fourier aus seinem Doppelintegral abgeleitet. (*Théor. anal. de la chaleur* No. 348.)

Ein weiteres Beispiel einer diskontinuierlichen Funktion bietet das Integral, das man aus den Integralen (7) ableiten kann:

$$(9) \int_0^\pi \log(1 - 2a\cos x + a^2)\, dx = 0, \quad \text{wenn } |a| \leqq 1,$$

$$= 2\pi \log |a|, \quad \text{wenn } |a| \geqq 1.$$

Für $a = \pm 1$ ist dies gleichbedeutend mit:

$$\int_0^{\frac{\pi}{2}} \log \sin x\, dx = \int_0^{\frac{\pi}{2}} \log \cos x\, dx = -\frac{\pi}{2} \log 2.$$

$$(10) \int_0^1 \frac{x^{a-1} - x^{-a}}{1-x} dx = \int_1^\infty \frac{x^{a-1} - x^{-a}}{1-x} dx$$

$$= \int_0^\infty \frac{x^{a-1} - x^{-\frac{1}{2}}}{1-x} dx = \pi \operatorname{ctg} a\pi,$$

$$\int_0^1 \frac{x^{a-1} + x^{-a}}{1+x} dx = \int_0^\infty \frac{x^{a-1}}{1+x} dx = \frac{\pi}{\sin a\pi}, \quad 0<a<1,$$

$$\int_0^1 \frac{x^a + x^{-a}}{1 + 2x\cos\theta + x^2} dx = \int_1^\infty \frac{x^a + x^{-a}}{1 + 2x\cos\theta + x^2} dx$$

$$= \int_0^\infty \frac{x^{-a}}{1 + 2x\cos\theta + x^2} dx = \frac{\pi \sin a\theta}{\sin a\pi \sin\theta}, \quad -\pi < \theta < \pi.$$

II. Durch *partielle Integration* werden z. B. die Integrale gefunden:

$$\int_0^{\frac{\pi}{2}} (\sin x)^m (\cos x)^{2n+1} dx = \frac{2 \cdot 4 \cdot 6 \cdots 2n}{(m+1)(m+3)\cdots(m+2n+1)},$$

$$\int_0^{\frac{\pi}{2}} (\sin x)^{2m+1} (\cos x)^{2n} dx = \frac{2 \cdot 4 \cdot 6 \cdots 2m \cdot 1 \cdot 3 \cdot 5 \cdots (2n-1)}{1 \cdot 3 \cdot 5 \cdots (2m+2n+1)},$$

$$\int_0^{\frac{\pi}{2}} (\sin x)^{2m} (\cos x)^{2n} dx = \frac{\pi}{2} \frac{1 \cdot 3 \cdot 5 \cdots (2m-1) \cdot 1 \cdot 3 \cdot 5 \cdots (2n-1)}{2 \cdot 4 \cdot 6 \cdots 2(m+n)}.$$

III. Die letzten Integrale gehören, ebenso wie die folgenden, der Theorie der *Gammafunktion* an, die in einem besonderen Kapitel behandelt wird.

$$\int_0^{\infty} e^{-x^2} dx = \frac{1}{2}\sqrt{\pi}.$$

Dieses wichtige Integral (Euler 1730) wird gewöhnlich nach Laplace (*Paris. Mém. de l'acad.* 1778) mit Hilfe von Doppelintegralen gefunden. Eine ganz direkte Ableitung auf Grund der Cauchy-Riemannschen Integraldefinition bei Cesàro, *Algebr. Analysis* 699. Eine andere nach Stieltjes, ebd. 719. Vgl. Hermite, *Cours*, 4. éd. 116.

$$\int_{-\infty}^{+\infty} e^{-ax^2-bx} dx = \sqrt{\frac{\pi}{a}} e^{\frac{b^2}{4a}}, \quad \int_{-\infty}^{+\infty} e^{-ax^2} \cos bx \, dx = \sqrt{\frac{\pi}{a}} e^{-\frac{b^2}{4a}},$$

$$\int_{-\infty}^{+\infty} e^{-ax^2-\frac{b}{x^2}} dx = \sqrt{\frac{\pi}{a}} \cdot e^{-2\sqrt{ab}},$$

$$\int_{-\infty}^{+\infty} \cos(x^2) dx = \int_{-\infty}^{+\infty} \sin(x^2) dx$$

$$= \int_0^{\infty} \frac{\cos x}{\sqrt{x}} dx = \int_0^{\infty} \frac{\sin x}{\sqrt{x}} dx = \sqrt{\frac{\pi}{2}}.$$

Die letzten Integrale (Euler 1781) sind spezielle Werte der *Fresnelschen Integrale*, die in der Beugungstheorie des Lichts

von Wichtigkeit sind. (Fresnel, *Œuvres* **1**, 319, Verdet, *Leç. d'optique physique* **1**, 328, Loria, *Spezielle ebene Kurven*, 457.)

$$a\int_0^\infty \frac{\cos bx\,dx}{a^2+x^2} = \varepsilon_a \frac{\pi}{2} e^{-|ab|}, \qquad \varepsilon_a \text{ Vorzeichen von } a.$$

$$\int_0^\infty \frac{x \sin bx}{a^2+x^2}\,dx = \varepsilon_b \frac{\pi}{2} e^{-|ab|}, \qquad \varepsilon_b \text{ Vorzeichen von } b.$$

IV. Durch *Reihenentwicklung* findet man z. B.

$$\int_0^1 \frac{\log x}{x-1}\,dx = \int_0^\infty \frac{x\,dx}{e^x-1} = \int_0^\infty x(e^{-x}+e^{-2x}+\cdots)\,dx =$$

$$= 1 + \frac{1}{2^2} + \frac{1}{3^2} + \cdots = \frac{\pi^2}{6}.$$

Ebenso

$$-\int_0^1 \frac{\log x}{x+1}\,dx = \int_0^\infty \frac{x\,dx}{e^x+1} = 1 - \frac{1}{2^2} + \frac{1}{3^2} - \cdots = \frac{\pi^2}{12},$$

$$\int_0^1 \frac{\log x}{x^2-1}\,dx = \int_0^\infty \frac{x\,dx}{e^x-e^{-x}} = 1 + \frac{1}{3^2} + \frac{1}{5^2} + \cdots = \frac{\pi^2}{8}.$$

Diese Integrale gehören der Theorie der *Riemannschen ζ-Funktion* an. (Vgl. Kap. XXII.)

Bei trigonometrischen Funktionen wird häufig ein unendlich großes Integrationsintervall in Teilintervalle von der Länge $\frac{\pi}{2}$ oder π zerlegt, die dann mit Hilfe der Periodizitätseigenschaften sämtlich auf das Intervall $\left(0, \frac{\pi}{2}\right)$ oder $(0, \pi)$ reduziert werden. So ist z. B.

$$\int_0^\infty \frac{\operatorname{tg} x}{x}\,dx = \int_0^{\frac{\pi}{2}} \operatorname{tg} x\left(\frac{1}{x} + \frac{1}{x-\pi} + \frac{1}{x+\pi} + \frac{1}{x-2\pi} + \frac{1}{x+2\pi} + \cdots\right)dx$$

$$= \int_0^{\frac{\pi}{2}} \operatorname{tg} x \cdot \operatorname{ctg} x\,dx = \frac{\pi}{2}.$$

$$\int_0^\infty \frac{\sin x}{1-2a\cos x+a^2}\,\frac{dx}{x}$$

$$=\int_0^\pi \frac{\sin x}{1-2a\cos x+a^2}\left(\frac{1}{x}+\frac{1}{x-2\pi}+\frac{1}{x+2\pi}+\frac{1}{x-4\pi}+\frac{1}{x+4\pi}+\cdots\right)dx$$

$$=\frac{1}{2}\int_0^\pi \frac{\sin x}{1-2a\cos x+a^2}\left(\operatorname{ctg} x+\frac{1}{\sin x}\right)dx$$

$$=\frac{\pi}{2(1-a)}, \quad |a|<1$$

$$=\frac{\pi}{2a(a-1)}, \quad |a|>1.$$

Fast alle hier genannten Integrale sind von Euler gefunden worden. Vgl. *Miscell. Berol.* **7**, 129 (1743), *Petersb. Akad.* 1777, Pars I, 3, *Opusc. anal.* **2**, 42 (1785) und namentlich *Inst. calc. int.* **4** (1794). Die wichtigsten Fortschritte über das von ihm Erreichte hinaus wurden durch den *Integralsatz von Cauchy* erzielt, der in der Theorie der analytischen Funktionen zu besprechen ist. Er bietet ein hervorragendes Hilfsmittel zur Ermittlung von bestimmten Integralen.

Wenn ein Integral sich nicht durch einen geschlossenen Ausdruck aus einer endlichen Anzahl von bekannten Funktionen darstellen läßt, so definiert es eine neue transzendente Funktion, zu deren Berechnung man sich gewöhnlich der *Reihenentwicklung* des Integrals bedient. Wichtige Beispiele von solchen Funktionen werden in besonderen Kapiteln behandelt werden (elliptische, hyperelliptische und Abelsche Integrale, Gammafunktion, Besselsche Funktionen).

§ 6. Kurvenintegrale.

Durch zwei reelle begrenzte Funktionen der reellen Variablen t:

$$x=\varphi(t), \quad y=\psi(t)$$

wird eine *Kurve* in der xy-Ebene definiert (die natürlich nicht notwendig *regulär* (Differentialrechnung § 2) zu sein braucht). Einem Intervall (a, b) der Veränderlichen t entspricht ein Kurvenstück C. Dieses Intervall wird in irgendeiner Weise in n Teilintervalle zerlegt mit den Endpunkten $a, t_1, t_2, \ldots, t_{n-1}, b$, denen die Funktionswerte $x_0, x_1, x_2, \ldots, x_{n-1}, x_n$ und $y_0, y_1,$

$y_2, \ldots, y_{n-1}, y_n$ entsprechen mögen. Zwischen diesen wählt man Werte $\xi_1, \xi_2, \ldots, \xi_n$ und $\eta_1, \eta_2, \ldots, \eta_n$ aus, so daß $x_{i-1} \leqq \xi_i \leqq x_i$, $y_{i-1} \leqq \eta_i \leqq y_i$ ist, und bildet für irgendeine begrenzte Funktion $P(x, y)$ von zwei Veränderlichen die Summe $\sum_{i=1}^{n} P(\xi_i, \eta_i)(x_i - x_{i-1})$. Wenn diese Summe bei unendlich wachsendem n, während die Teilintervalle $(t_{i-1} t_i)$ sämtlich gegen Null konvergieren, einen bestimmten endlichen Grenzwert besitzt, so heißt die Funktion $P(x, y)$ *längs C integrierbar*, und der Grenzwert heißt *das in bezug auf x längs C erstreckte Kurvenintegral von* $P(x, y)$. Man schreibt es

$$\int_C P(x, y)\, dx = \lim_{n=\infty} \sum P(\xi_i, \eta_i)(x_i - x_{i-1}).$$

Entsprechend wird das in bezug auf y längs C erstreckte Kurvenintegral einer Funktion $Q(x, y)$ definiert durch

$$\int_C Q(x, y)\, dy = \lim_{n=\infty} \sum Q(\xi_i, \eta_i)(y_i - y_{i-1}),$$

und man schreibt

$$\int_C P(x, y)\, dx + \int_C Q(x, y)\, dy = \int_C (P\, dx + Q\, dy).$$

Die Existenz des Kurvenintegrals hängt von der Beschaffenheit der Kurve, also von den Funktionen $\varphi(t)$ und $\psi(t)$, sowie von der zu integrierenden Funktion ab. Es gilt darüber folgendes:

Sind $\varphi(t)$ *und* $\psi(t)$ *im Intervall* (a, b) *eindeutige, stetige Funktionen von beschränkter Schwankung* (S. 39) *und bedeutet* γ_i *die Maximalschwankung von* $P(x, y)$ *in der Umgebung*[1]) *des Punktes* (ξ_i, η_i), *so ist zur Integrierbarkeit von* $P(x, y)$ *längs C in bezug auf x notwendig und hinreichend, daß* $\lim \sum \gamma_i |x_i - x_{i-1}| = 0$ *ist.*

Die hier gegebene, durch große Allgemeinheit ausgezeichnete Definition der Kurvenintegrale rührt von Heffter, *Gött. Nachr.* (1902), 115, her. Vgl. Picard, *Traité d'analyse* 1. éd. **1**, 70, Jordan, *Cours d'analyse* 2. éd. **1**, 181, Pringsheim, *Münch. Ber.* **25**, 39 u. 295 (1895).

Infolge der Voraussetzungen über die Funktionen $\varphi(t)$ und $\psi(t)$ ist der *Integrationsweg C* eine *stetige rektifizierbare Kurve*

1) Über die genauere Begrenzung dieser Umgebung vgl. Heffter, *Gött. Nachr.* (1902), 120.

im Sinne von C. Jordan (vgl. auch Study, *Math. Ann.* **47**, 313 (1896), u. Schoenflies, *Entwickl. d. Lehre v. d. Punktmannigf.* **2**, 246 (1908).

Unter engeren, aber immer noch sehr allgemeinen Annahmen über die Beschaffenheit des Integrationsweges kann man *jedes Kurvenintegral auf ein gewöhnliches bestimmtes Integral zurückführen*, nämlich:

a) Sind $x = \varphi(t)$ und $y = \psi(t)$ im Intervall (a, b) *monotone* Funktionen und ergeben sich durch Elimination von t die Gleichungen $y = f(x)$, $x = g(y)$, so ist

$$\int_C P(x, y)\,dx = \int_{\varphi(a)}^{\varphi(b)} P(x, f(x))\,dx; \qquad \int_C Q(x, y)\,dy = \int_{\psi(a)}^{\psi(b)} Q(g(y), y)\,dy.$$

Sind $\varphi(t)$ und $\psi(t)$ *abteilungsweise monoton*, so lassen sich die Kurvenintegrale aus einer endlichen Anzahl solcher gewöhnlichen Integrale zusammensetzen.

b) Besitzen $\varphi(t)$ und $\psi(t)$ im Intervall (a, b) *stetige Ableitungen*, so ist

$$\int_C P(x, y)\,dx = \int_a^b P(\varphi(t), \psi(t))\,\varphi'(t)\,dt;$$

$$\int_C Q(x, y)\,dy = \int_a^b Q(\varphi(t), \psi(t))\,\psi'(t)\,dt.$$

1. *Fundamentalsatz der Integralrechnung für Kurvenintegrale.* Ist eine Funktion $F(x, y)$ *nebst ihren ersten Ableitungen* $\frac{\partial F}{\partial x}$ und $\frac{\partial F}{\partial y}$ längs der Kurve C und in ihrer Umgebung *eindeutig* und *stetig* und sind $F(\!(a)\!)$ und $F(\!(b)\!)$ die Funktionswerte in den Endpunkten des Intervalls (a, b), so ist

$$\int_C dF = \int_C \left(\frac{\partial F}{\partial x}\,dx + \frac{\partial F}{\partial y}\,dy\right) = F(\!(b)\!) - F(\!(a)\!).$$

Insbesondere ist längs einer *geschlossenen* Kurve $\int_C dF = 0$, und wenn die Funktion F und ihre ersten Ableitungen in einem zusammenhängenden *Bereich* eindeutig und stetig sind, so ist ein zwischen zwei festen Punkten des Bereichs genommenes Kurvenintegral $\int dF$ *unabhängig vom Integrationsweg*, solange dieser den Bereich nirgends verläßt.

2. *Sind $P(x, y)$ in einem zusammenhängenden Bereich eindeutig und stetig und ist das von einem festen Punkt (x_0, y_0) bis zum Punkt (x, y) erstreckte Integral* $\int_C (P\,dx + Q\,dy) = F(x, y)$ *in diesem Bereich unabhängig vom Weg, so ist F dort eine stetige Funktion von x, y, und es ist* $P(x, y) = \frac{\partial F}{\partial x}$, $Q(x, y) = \frac{\partial F}{\partial y}$, *also $P\,dx + Q\,dy$ das vollständige Differential von F.*

Von größter Bedeutung ist die Umkehrung dieses Satzes:

3a. *Es seien die Funktionen $P(x, y)$ und $Q(x, y)$, sowie die Ableitungen* $\frac{\partial P}{\partial y}$ *und* $\frac{\partial Q}{\partial x}$ *in einem Bereiche eindeutig und stetig; damit dann das zwischen zwei Punkten des Bereichs genomm Kurvenintegral* $\int_C (P\,dx + Q\,dy)$ *vom Integrationsweg unabhängig ist, solange dieser den Bereich nirgends verläßt, ist notwendig und hinreichend, daß*

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$$

ist.

Dieser Satz, der sich schon bei Clairaut, *Théorie de la figure de la terre* (1743) findet (vgl. Stäckel, *Bibl. math.* (8) **2**, 111 (1901)), bildet die Grundlage des Cauchyschen Satzes in der Lehre von den analytischen Funktionen und ist unter den angegebenen Voraussetzungen von Cauchy, *C. R.* **23**, 251 (1846), bewiesen worden. Es hat aber Goursat, *Amer. Trans.* **1**, 14 (1900), gezeigt, daß es nicht notwendig ist, die *Stetigkeit* der Ableitungen $\frac{\partial P}{\partial y}$ und $\frac{\partial Q}{\partial x}$, sondern nur ihre *Existenz* vorauszusetzen. Hieran knüpfen die neueren Untersuchungen über diesen Satz an, von denen Heffter, *Gött. Nachr.* (1903), 312 und Pringsheim, *Münch. Ber.* **33** (1903), 674 als abschließend bezeichnet werden können.

Gewöhnlich wird der Satz in folgender Form ausgesprochen:

3b. *Unter den angegebenen Voraussetzungen ist die Bedingung* $\frac{\partial P}{\partial y} - \frac{\partial Q}{\partial x} = 0$ *notwendig und hinreichend, damit das über irgendeine geschlossene Kurve des Bereichs genommene Integral* $\int_C (P\,dx + Q\,dy) = 0$ *ist.*

§ 7. Mehrfache Integrale.

Ein durch eine reguläre geschlossene Kurve begrenztes Gebiet T der xy-Ebene werde irgendwie in n Teilgebiete $d_1, d_2 \ldots d_n$ zerlegt und in jedem Teilgebiet d_i werde ein Punkt (x_i, y_i) angenommen. Bildet man dann für eine im ganzen Gebiet stetige Funktion $f(x, y)$ die Summe $\sum_{i=1}^{n} d_i f(x_i, y_i)$ und vermehrt die Anzahl n der Teilgebiete unbegrenzt, während diese gleichzeitig nach allen Seiten hin unendlich klein werden, so besitzt die Summe einen bestimmten, von der Art der Einteilung und der Wahl der Punkte (x_i, y_i) unabhängigen Grenzwert, den man als *das über das Gebiet T erstreckte Doppelintegral*

$$\iint_T f(x, y)\, dx\, dy$$

bezeichnet. Die Schreibweise rührt her von einer Einteilung des Gebiets in *rechteckige* Elemente durch zwei Systeme von Parallelen zu den Koordinatenachsen. Vgl. hierzu die sehr anschauliche Darstellung von Weber, *Arch. Math. u. Phys.* (3) **15**, 289 (1909).

Man kann die Definition des Doppelintegrals auch auf den Fall ausdehnen, daß die Funktion $f(x, y)$ innerhalb T linienweise unstetig oder in einzelnen Punkten unendlich wird oder daß der Bereich T sich ins Unendliche erstreckt. Über die strenge Begründung der Lehre von den Doppelintegralen vgl. Arzelà, *Bologna Mem.* (5) **2**, 133 (1892), de la Vallée-Poussin, *J. d. math.* (4) **8**, 400 (1892), ebd. (5) **5**, 202 (1899), Jordan, *Cours d'anal.* **1**, (1893), Pringsheim, *Münch. Ber.* **28**, 69 (1898), ebd. **29**, 47 (1899), Stolz, *Grundzüge* **3** (1899), Freud, *Monatsh. f. Math.* **18**, 29 (1907).

Geometrisch bedeutet $z = f(x, y)$ eine *Fläche*, und das Doppelintegral gibt das *Volumen* des geraden Zylinders, der durch diese Fläche und durch das Gebiet T in der xy-Ebene begrenzt ist. Denkt man sich T so mit *Masse* belegt, daß $f(x, y)$ die *Dichtigkeit* an der Stelle (x, y) angibt, so bedeutet das Doppelintegral die *Gesamtmasse* der Belegung von T.

Das Integrationsgebiet T kann durch eine Ungleichung von der Form $\varphi(x, y) \leqq 0$ charakterisiert werden, und wir können annehmen, daß die Begrenzungslinie $\varphi(x, y) = 0$ durch jede Parallele zu den Achsen höchstens in zwei Punkten getroffen wird (andernfalls wird man T in eine endliche Anzahl von derartigen Bereichen zerlegen können); dann wird die Kurve von

einer Parallelen mit der Abszisse x in zwei Punkten mit den Ordinaten $y_1 = \varphi_1(x)$ und $y_2 = \varphi_2(x)$ und von einer Parallelen mit der Ordinate y in zwei Punkten mit den Abszissen $x_1 = \psi_1(y)$ und $x_2 = \psi_2(y)$ geschnitten, und wenn x zwischen a und b, y zwischen c und d variieren kann, *so kann man das Doppelintegral auf zweifache Weise durch zwei aufeinanderfolgende einfache Integrationen ausdrücken*, nämlich

$$\iint_T f(x,y)\,dx\,dy = \int_a^b \Big[\int_{\varphi_1(x)}^{\varphi_2(x)} f(x,y)\,dy\Big]dx = \int_c^d \Big[\int_{\psi_1(y)}^{\psi_2(y)} f(x,y)\,dx\Big]dy.$$

Als besonderer Fall ist hierin der Satz von der *Vertauschbarkeit der Integrationen* enthalten, wenn nämlich T ein *Rechteck* ist, dessen Seiten zu den Achsen parallel sind, also überall $\varphi_1(x) = c$, $\varphi_2(x) = d$, $\psi_1(y) = a$, $\psi_2(y) = b$. Nimmt man als Bereich T ein gleichschenklig rechtwinkliges Dreieck mit den Seiten $y = 0$, $x = a$, $y = x$, so erhält man eine Formel von Dirichlet

$$\int_0^a \Big[\int_0^x f(x,y)\,dy\Big]dx = \int_0^a \Big[\int_y^a f(x,y)\,dx\Big]dy.$$

Ebenso wie über einen *ebenen* Bereich kann man auch Integrale definieren, die über ein Gebiet T einer *krummen Oberfläche* erstreckt sind. Ist f eine Funktion des Ortes auf der Fläche, do das Oberflächenelement, so kann ein solches *Oberflächenintegral* durch $\int_T f\,do$ bezeichnet werden. Die Oberfläche kann durch drei Gleichungen mit zwei Parametern

$$x = x(u,v), \quad y = y(u,v), \quad z = z(u,v)$$

dargestellt werden und der Bereich T ist dann durch eine Ungleichung von der Form $\varphi(u,v) \leqq 0$ gegeben. Nach den Grundformeln der Flächentheorie ist das Oberflächenelement $do = \sqrt{EG - F^2}\,du\,dv$, worin E, F, G die *Fundamentalgröße erster Ordnung*

$$E = \sum_{x,y,z}\Big(\frac{\partial x}{\partial u}\Big)^2, \quad F = \sum_{x,y,z}\frac{\partial x}{\partial u}\,\frac{\partial x}{\partial v}, \quad G = \sum_{x,y,z}\Big(\frac{\partial x}{\partial v}\Big)^2$$

bedeuten, und es wird damit das Oberflächenintegral auf ein Doppelintegral:

$$\int_T f\,do = \iint f(u, v)\sqrt{EG - F^2}\,du\,dv$$

zurückgeführt.

Entsprechend wie bei den Doppelintegralen verfährt man, um *vielfache Integrale* zu definieren. So stellt sich das *dreifache Integral*

$$\iiint_V f(x, y, z)\,dx\,dy\,dz = \int_V f\,dv$$

dar als Grenzwert einer Summe, erstreckt über die Elemente dv eines allseitig begrenzten Raumteils V, jedes Element multipliziert mit dem Wert der Funktion $f(x, y, z)$ in einem zum Element gehörigen Punkt. Die Auswertung eines solchen Integrals kann durch drei aufeinanderfolgende einfache Integrationen geschehen. Für ein *n-faches Integral*

$$\int\int\cdots\int_V f(x_1, x_2, \ldots, x_n)\,dx_1\,dx_2\ldots dx_n$$

wird ein endlicher Bereich V einer n-dimensionalen Mannigfaltigkeit durch eine Ungleichung $\varphi(x_1, x_2, \ldots x_n) \leqq 0$ abgegrenzt und in entsprechender Weise über sämtliche Elemente des Bereichs summiert.

Einführung von neuen Variablen in mehrfache Integrale. Treten an Stelle der Variablen $x_1, x_2, \ldots, x_n$ in einem n-fachen Integral neue durch die Gleichungen $x_i = \varphi_i(y_1, y_2, \ldots, y_n)^{(i=1,2,\ldots,n)}$, so daß jedem Wertesystem $(x_1, x_2, \ldots, x_n)$ ein bestimmtes Wertesystem $(y_1, y_2, \ldots, y_n)$ eindeutig zugeordnet ist und umgekehrt, so ist

$$\int\int\cdots\int f(x_1\ldots x_n)\,dx_1\ldots dx_n = \int\int\cdots\int f(\varphi_1\ldots\varphi_n)\,|D|\,dy_1\ldots dy_n,$$

und darin bedeutet D die *Funktionaldeterminante* (vgl. S. 156):

$$D = \frac{\partial(x_1\,x_2\ldots x_n)}{\partial(y_1\,y_2\ldots y_n)}.$$

Sie muß, damit die Variablen $y_1 \ldots y_n$ voneinander unabhängig sind, von Null verschieden sein. Für Doppelintegrale bereits bei Euler, *Novi Comm. Petrop.* **14**, 72 (1769), für dreifache Integrale Lagrange, *Nouv. Mém.* Berlin (1773), 125, allgemein Jacobi, *J. f. Math.* **12**, 38 (1833).

Von größter Bedeutung, namentlich für Funktionentheorie und mathematische Physik sind die Sätze, durch welche es möglich wird, an Stelle von Integrationen über einen ganzen Bereich

nur solche über die *Begrenzung* auszuführen. (Gauß, *Theoria attract. corp. sphaeroid.* (1813), § 9.)

1. *Satz von Gauß.* (Allgem. Lehrs. über . . . Anziehungskr. (1840), § 24. *Ostwalds Klassiker* No. 2.)

Es seien A, B, C innerhalb eines endlichen Raumes V eindeutige stetige Funktionen von x, y, z; der Raum V sei begrenzt von der Oberfläche T mit dem Flächenelement do und $\cos\alpha$, $\cos\beta$, $\cos\gamma$ seien die Richtungskosinus der ins Innere von V gerichteten Flächennormale; dann ist

$$(1) \qquad \int_V \left(\frac{\partial A}{\partial x} + \frac{\partial B}{\partial y} + \frac{\partial C}{\partial z}\right) dv = -\int_T (A\cos\alpha + B\cos\beta + C\cos\gamma)\, do.$$

Auf die Ebene übertragen gibt dieser Satz die Transformation eines Integrals über eine *ebene Fläche* F in ein Integral über die *Randkurve* C, nämlich für zwei Funktionen $P(x, y)$ und $Q(x, y)$:

$$\iint_F \left(\frac{\partial P}{\partial y} - \frac{\partial Q}{\partial x}\right) dx\, dy = -\int_C (P\, dx + Q\, dy).$$

Auf diesem Satz beruhen die älteren Beweise für die Sätze 1, 2 und 3 in § 6.

Es ist nach (1) das über die *geschlossene* Oberfläche T genommene Integral auf der rechten Seite *stets gleich Null, sobald im ganzen Innern von* V

$$(2) \qquad \frac{\partial A}{\partial x} + \frac{\partial B}{\partial y} + \frac{\partial C}{\partial z} = 0$$

ist, und diese Bedingung ist notwendig und hinreichend, damit das über *irgendein Flächenstück* F *im Innern von* V erstreckte Integral nur abhängig ist von der *Randkurve* C des Stückes. (Übertragung von Satz 3 in § 6 auf den Raum. Picard, *Traité d'analyse* **1**, 114.) Dies kommt zum Ausdruck durch den

2. *Satz von Stokes.* (*Cambridge Univ. calend.* 1854.)

Wird

$$A = \frac{\partial Q}{\partial z} - \frac{\partial R}{\partial y}, \quad B = \frac{\partial R}{\partial x} - \frac{\partial P}{\partial z}, \quad C = \frac{\partial P}{\partial y} - \frac{\partial Q}{\partial x}$$

gesetzt, so ist die Bedingung (2) erfüllt, und umgekehrt lassen sich bei Bestehen der Gleichung (2) immer drei derartige Funktionen P, Q, R finden. Es ist dann

$$\int_F \left[\left(\frac{\partial Q}{\partial z}-\frac{\partial R}{\partial y}\right)\cos\alpha+\left(\frac{\partial R}{\partial x}-\frac{\partial P}{\partial z}\right)\cos\beta+\left(\frac{\partial P}{\partial y}-\frac{\partial Q}{\partial x}\right)\cos\gamma\right]do$$

$$(3)\qquad =-\int_C (P\,dx+Q\,dy+R\,dz).$$

3. *Satz von Green* (*J. f. Math.* **44**, 360 (1828)). Setzt man in (1)

$$A=\varphi\frac{\partial\psi}{\partial x},\qquad B=\varphi\frac{\partial\psi}{\partial y},\qquad C=\varphi\frac{\partial\psi}{\partial z},$$

so erhält man

$$\int_V \left(\frac{\partial\varphi}{\partial x}\frac{\partial\psi}{\partial x}+\frac{\partial\varphi}{\partial y}\frac{\partial\psi}{\partial y}+\frac{\partial\varphi}{\partial z}\frac{\partial\psi}{\partial z}\right)dv+\int_V \varphi\,\Delta\psi\,dv$$

$$(4)\qquad =-\int_T \varphi\left[\frac{\partial\psi}{\partial x}\cos\alpha+\frac{\partial\psi}{\partial y}\cos\beta+\frac{\partial\psi}{\partial z}\cos\gamma\right]do$$

$$=-\int_T \varphi\frac{d\psi}{dn}do,$$

wobei $\Delta\psi$ den *Differentialparameter* $\frac{\partial^2\psi}{\partial x^2}+\frac{\partial^2\psi}{\partial y^2}+\frac{\partial^2\psi}{\partial z^2}$ und $\frac{d\psi}{dn}$ den Differentialquotienten längs der ins Innere von V gerichteten Flächennormale bedeutet. Durch Vertauschung von φ und ψ und Subtraktion folgt daraus der Satz:

$$(5)\qquad \int_V (\varphi\Delta\psi-\psi\Delta\varphi)\,dv=-\int_T\left(\varphi\frac{d\psi}{dn}-\psi\frac{d\varphi}{dn}\right)do.$$

Im weiteren Sinne bezeichnet man als Greensche Sätze Formeln vom gleichen Charakter, wie (4) und (5), in denen an Stelle des Differentialparameters allgemeinere Differentialausdrücke 2. Ordnung auftreten. Vgl. die Ausführungen über *Randwertaufgaben* in Kap. XIV.

Kapitel IX.

Differenzenrechnung.

Von *H. E. Timerding* in Braunschweig.

§ 1. Differenzen.

Die Differenzenrechnung ist die Zwillingsschwester der Differentialrechnung. Die Entdecker der letzteren, Leibniz und Newton, haben gleichzeitig auch zuerst das Wesen und den methodischen Charakter der ersteren erkannt. (Vgl. Cantor, *Gesch. d. Math.* **3** (1898) S. 37, 358 und Zeuthen, *Gesch. d. Math. im XVI. u. XVII. Jahrh.*) Als eigentlicher Begründer der Differenzenrechnung hat Newton zu gelten, der zu ihrer Entwicklung in einer kleinen, 1711 erschienenen Schrift *Methodus differentialis* den Grund legte. (Diese Schrift ist wiederabgedruckt in den *Opuscula math.* von Newton 1744 und steht in der Ausgabe von Horsley (1779) in Bd. 1, S. 519.) Als das erste Lehrbuch ist anzusehen Brook Taylor's *Methodus incrementorum directa et inversa,* London 1715, dem 1730 Stirlings *Methodus differentialis* folgte. Eine weitere methodische Ausgestaltung hat die Differenzenrechnung in Eulers *Institutiones calculi differentialis,* Petersburg 1755, erfahren.

Die Differenzenrechnung setzt eine tabellarisch gegebene Funktion voraus, d. h. man nimmt an, die Werte der Funktion $f(x)$ seien für eine Reihe von Werten des Argumentes x, die in den gleichen Abständen h aufeinander folgen, bekannt. Dann schreibt man:

$$\Delta f(x) = f(x+h) - f(x)$$

und bezeichnet dies als die erste Differenz der Funktion $f(x)$. Verfährt man mit der ersten Differenz ebenso wie vorher mit der Funktion selbst, so erhält man die zweite Differenz $\Delta^2 f(x)$ usw. Allgemein ist:

$$\Delta^n f(x) = \Delta^{n-1} f(x+h) - \Delta^{n-1} f(x),$$

und man findet:

$$\Delta^{m+n} f(x) = \Delta^m \Delta^n f(x).$$

Bezeichnet man der Kürze halber die Werte

$$x_0 \quad x_0 + h \quad x_0 + 2h \ldots x_0 + nh$$

mit

$$x_0 \quad x_1 \quad x_2 \quad \ldots \quad x_n,$$

so wird:

$$\Delta^n f(x_0) = f(x_n) - \binom{n}{1} f(x_{n-1}) + \binom{n}{2} f(x_{n-2}) - \cdots \pm f(x_0).$$

Mit Hilfe der sukzessiven Differenzen für den Wert des Argumentes x_0 kann man den Funktionswert für x_n folgendermaßen ausdrücken:

$$f(x_n) = f(x_0) + \binom{n}{1} \Delta f(x_0) + \binom{n}{2} \Delta^2 f(x_0) + \cdots + \Delta^n f(x_0)$$

oder symbolisch:

$$f(x_n) = (1 + \Delta)^n f(x_0).$$

Die Differenz einer ganzen Funktion von x ist eine ganze Funktion, deren Grad um 1 niedriger ist. Die n^{te} Differenz einer ganzen Funktion n^{ten} Grades ist somit eine Konstante, und zwar wird, wenn

$$f(x) = a_0 x^n + a_1 x^{n-1} + \cdots$$

angenommen wird:

$$\Delta^n f(x) = n!\, a_0 \cdot h^n.$$

Insbesondere ergibt sich:

$$\Delta x(x-h)(x-2h)\ldots(x-nh) = (n+1)h \cdot x(x-h)\ldots(x-(n-1)h),$$

außerdem die von Nicole (*Hist. de l'Acad. des Sciences de Paris*, Année 1717, p. 7) gefundene Formel:

$$\Delta \frac{1}{x(x+h)(x+2h)\ldots(x+(n-1)h)} = -nh \frac{1}{x(x+h)\ldots(x+nh)}.$$

Die Differenzen der Exponentialfunktion sind dem Funktionswerte proportional. Es wird:

$$\Delta^n a^x = (a^h - 1)^n \cdot a^x.$$

Für die Kreisfunktionen findet man:

$$\Delta^n \sin u = \left(2 \sin \frac{1}{2} h\right)^n \cdot \sin\left(u + \frac{n}{2}(\pi + h)\right),$$

$$\Delta^n \cos u = \left(2 \sin \frac{1}{2} h\right)^n \cdot \cos\left(u + \frac{n}{2}(\pi + h)\right).$$

Es gelten die allgemeinen Regeln:

$$\Delta c f(x) = c \Delta f(x) \quad (c = \text{const.}),$$

$$\Delta[\varphi(x) + \psi(x)] = \Delta\varphi(x) + \Delta\psi(x),$$

$$\Delta[\varphi(x) \cdot \psi(x)] = \varphi(x) \cdot \Delta\psi(x) + \psi(x+h) \cdot \Delta\varphi(x).$$

Allgemein wird für $x + nh = x_n$:

$$\Delta^n[\varphi(x) \cdot \psi(x)] = \varphi(x_n) \cdot \Delta^n\psi(x) + \binom{n}{1} \Delta\varphi(x_n - h) \cdot \Delta^{n-1}\psi(x)$$
$$+ \binom{n}{2} \Delta^2\varphi(x_n - 2h) \cdot \Delta^{n-2}\psi(x) + \cdots + \Delta^n\varphi(x) \cdot \psi(x).$$

Die Differenzen werden zu den Differentialquotienten in Beziehung gesetzt durch folgende Formel:

$$\Delta^m f(x) = A_m^m h^m f^{(m)}(x) + \cdots + A_m^n h^n f^{(n)}(x) + A_m^{n+1} h^{n+1} f^{(n+1)}(\xi),$$

wo ξ einen Wert zwischen x und $x + nh$ bedeutet, ferner

$$f^{(\mu)}x = \frac{d^\mu f(x)}{dx^\mu}$$

und (für $\mu \geqq m$)

$$A_m^\mu = \frac{1}{\mu!} \cdot \left\{ \frac{\Delta^m (x-a)^\mu}{h^\mu} \right\}_{x=a}^{h=1} = \frac{1}{\mu!} \left\{ \Delta^m t^\mu \right\}_{t=0}$$
$$= \frac{1}{\mu!} \left[m^\mu - \binom{m}{1} (m-1)^\mu + \binom{m}{2} (m-2)^\mu \cdots \pm m \right]$$

gesetzt ist. Umgekehrt wird:

$$h^m f^{(m)}(x) = B_m^m \Delta^m f(x) + \cdots + B_m^n \Delta^n f(x) + B_m^{n+1} h^{n+1} f^{(n+1)}(\xi).$$

Hierin ist ξ wieder ein Wert zwischen x und $x + nh$, und es ist:

$$B_m^\mu = \left\{ \frac{d^m}{dt^m} \left(\frac{t(t-1) \ldots (t-\mu+1)}{1 \cdot 2 \ldots \mu} \right) \right\}_{t=0}.$$

Die Koeffizienten A_m^μ und B_m^μ kann man auch durch die Reihenentwicklungen festlegen:

$$(e^h - 1)^m = A_m^m h^m + A_m^{m+1} h^{m+1} + \cdots,$$
$$[\log(1+t)]^m = B_m^m t^m + B_m^{m+1} t^{m+1} + \cdots.$$

Daraus folgt die nachstehende symbolische Schreibweise der obigen Gleichungen:

$$\Delta^m f(x) = \{ e^{hD} - 1 \}^m f(x), \quad \frac{d^m f(x)}{dx^m} = \left\{ \frac{\log(1+\Delta)}{h} \right\}^m f(x).$$

Die Gleichungen sind so zu verstehen, daß der symbolische Faktor von $f(x)$ auf der rechten Seite nach Potenzen von D, resp. Δ entwickelt und hierauf

$$D^\mu \cdot f(x) = \frac{d^\mu f(x)}{dx^\mu}, \quad \Delta^\mu \cdot f(x) = \Delta^\mu f(x)$$

gesetzt werden muß.

Mit Hilfe dieser symbolischen Schreibweise kann man leicht eine Formel geben, die es gestattet, von den für ein Intervall h gebildeten Differenzen $\Delta^\mu f(x)$ zu den für ein anderes Intervall h_1 gebildeten Differenzen $\Delta_1^\mu f(x)$ überzugehen. Diese Formel lautet:

$$\Delta_1^m f(x) = [(1 + \Delta)^{\frac{h_1}{h}} - 1]^m f(x).$$

§ 2. Interpolation.

Die Aufgabe der rationalen Interpolation verlangt zunächst, eine rationale, insbesondere eine ganze Funktion $f(x)$ so zu bestimmen, daß sie für gewisse Werte des Argumentes

$$x_0\, x_1\, x_2 \ldots x_n$$

vorgegebene Werte

$$y_0\, y_1\, y_2 \cdots y_n$$

annimmt. Es sind zwei verschiedene Lösungsformen dieser Aufgabe zu unterscheiden, die Lagrangesche und die Newtonsche

Die Lagrangesche Formel erhält man, wenn man

$$\varphi(x) = (x - x_0)(x - x_1) \ldots (x - x_n),$$

$$\varphi'(x_\mu) = \left[\frac{d\varphi(x)}{dx}\right]_{x = x_\mu} = (x_\mu - x_0) \ldots (x_\mu - x_{\mu-1}) \cdot (x_\mu - x_{\mu+1}) \ldots (x_\mu - x_n)$$

setzt, in der Gestalt:

$$f(x) = \sum_{\mu=0}^{n}{}' \frac{y_\mu}{\varphi'(x_\mu)} \cdot \frac{\varphi(x)}{x - x_\mu}.$$

Die Newtonsche Formel lautet in der Fassung, die ihr Gauß gegeben hat:

$$f(x) = y_0 + (x - x_0)[x_0 x_1] + (x - x_0)(x - x_1)[x_0 x_1 x_2] + (x - x_0)(x - x_1)(x - x_2)[x_0 x_1 x_2 x_3] + \cdots$$

Hierbei ist:

$$[x_0 x_1] = \frac{y_1 - y_0}{x_1 - x_0}, \quad [x_1 x_2] = \frac{y_2 - y_1}{x_2 - x_1}, \quad [x_2 x_3] = \frac{y_3 - y_2}{x_3 - x_2} \ldots$$

$$[x_0 x_1 x_2] = \frac{[x_1 x_2] - [x_0 x_1]}{x_2 - x_0}, \quad [x_1 x_2 x_3] = \frac{[x_2 x_3] - [x_1 x_2]}{x_3 - x_1} \ldots$$

$$[x_0 x_1 x_2 x_3] = \frac{[x_1 x_2 x_3] - [x_0 x_1 x_2]}{x_3 - x_0} \ldots$$

$$\ldots \ldots \ldots$$

Die Newtonsche Formel hat den Vorzug, daß man bei Hinzunahme neuer gegebener Größen nur neue Glieder hinzuzufügen braucht, ohne die vorhergehenden ändern zu müssen.

Die weitere Aufgabe ist nun die, von einer bestimmten Funktion $F(x)$, deren Werte für eine Reihe von Argumentenwerten gegeben sind, die Werte für alle Argumentenwerte eines gewissen Bereiches näherungsweise zu ermitteln. Die vorstehend angegebenen ganzen Funktionen $f(x)$ sind dann als Näherungsformeln für die Funktion $F(x)$ anzusehen, und es wird somit:

$$F(x) = f(x) + R,$$

wo R ein Korrektionsglied bedeutet. Nach Cauchy, *C. R.* **11** (1840) 787, *Œuvres* (1) **5**, 422, hat man zu setzen:

$$R = \frac{(x - x_0)(x - x_1) \ldots (x - x_n)}{(n+1)!} f^{(n+1)}(\xi),$$

wenn ξ einen Wert zwischen x_0 und x_n bedeutet.

Für den besonders wichtigen Fall, daß die Werte der Argumente, für welche die Funktionswerte gegeben sind, in gleichen Intervallen h aufeinander folgen, also von der Form sind:

$$x_0, x_0 + h, x_0 + 2h \ldots x_0 + nh,$$

geht die Newtonsche Formel über in:

$$f(x) = f(x_0) + \frac{x - x_0}{1!} \cdot \frac{\Delta f(x_0)}{h} + \frac{(x - x_0)(x - x_0 - h)}{2!} \cdot \frac{\Delta^2 f(x_0)}{h^2} + \ldots$$
$$+ \frac{(x - x_0)(x - x_0 - h) \ldots (x - x_0 - (n-1)h)}{n!} \cdot \frac{\Delta^n f(x_0)}{h^n}$$

Substituiert man:

$$x = x_0 + ht$$

und setzt:

$$f(x) = f(x_0 + ht) = \varphi(t), \quad f(x_0) = \varphi(0) = y_0,$$

so wird einfacher:

$$\varphi(t) = y_0 + t\Delta y_0 + \frac{t(t-1)}{2!} \Delta^2 y_0 + \frac{t(t-1)(t-2)}{3!} \Delta^3 y_0 + \cdots.$$

Diese Funktion nimmt für

$$t = 0, 1, 2, 3 \ldots$$

die verlangten Werte an, und so wird die Newtonsche Formel gewöhnlich gegeben.

Durch größere Symmetrie ausgezeichnet ist die Formel, die dazu dient, die Werte der Funktion in der Umgebung einer bestimmten Stelle zu bestimmen und die Stirlingsche Formel genannt wird, obwohl sie in einem von Newton gegebenen Satze (*Methodus differentialis*, Prop. III) bereits enthalten ist. Wir nehmen auch die Werte des Argumentes $t = -1, -2 \ldots$ hinzu und stellen die folgende Tabelle auf:

Arg.	Fkt.	Differenzen					
t	y	I	II	III	IV	V	VI
-3	$\bar{y}_3$						
		$\Delta\bar{y}_3$					
-2	$\bar{y}_2$		$\Delta^2\bar{y}_3$				
		$\Delta\bar{y}_2$		$\Delta^3\bar{y}_3$			
-1	$\bar{y}_1$		$\Delta^2\bar{y}_2$		$\Delta^4\bar{y}_3$		
		$\Delta\bar{y}_1$		$\Delta^3\bar{y}_2$		$\Delta^5\bar{y}_3$	
0	y_0		$\Delta^2\bar{y}_1$		$\Delta^4\bar{y}_2$		$\Delta^6\bar{y}_3$
		Δy_0		$\Delta^3\bar{y}_1$		$\Delta^5\bar{y}_2$	
1	y_1		$\Delta^2 y_0$		$\Delta^4\bar{y}_1$		
		Δy_1		$\Delta^3 y_0$			
2	y_2		$\Delta^2 y_1$				
		Δy_2					
3	y_3						

Setzen wir dann noch in leicht verständlicher Symbolik:

$$\tfrac{1}{2}(\Delta\bar{y}_1 + \Delta y_0) = \Delta\bar{y}_{\frac{1}{2}}, \quad \tfrac{1}{2}(\Delta^3\bar{y}_2 + \Delta^3\bar{y}_1) = \Delta^3\bar{y}_{\frac{3}{2}} \ldots,$$

so wird die Stirlingsche Formel, indem τ einen Wert zwischen $-\frac{1}{2}$ und $+\frac{1}{2}$ bedeutet:

$$\varphi(\tau) = y_0 + \tau\Delta\bar{y}_{\frac{1}{2}} + \frac{\tau^2}{2!}\Delta^2\bar{y}_1 + \frac{(\tau+1)\tau(\tau-1)}{3!}\Delta^3\bar{y}_{\frac{3}{2}}$$
$$+ \frac{(\tau+1)\tau\cdot\tau(\tau-1)}{4!}\Delta^4\bar{y}_2 + \frac{(\tau+2)(\tau+1)\tau(\tau-1)(\tau-2)}{5!}\Delta^5\bar{y}_{\frac{5}{2}} \cdots$$

Die wichtigste Literatur über Interpolation sei im folgenden zusammengestellt: Lagrange, *Sur les interpolations, Œuvres* **7**, 535; *Sur une méthode partic. d'approximation, Œuvres* **5**, 517; *Mém. sur la méthode d'interpolation, Œuvres* **5**, 663; *Sur une nouvelle espèce de calcul, Œuvres* **3**, 441. Gauß, *Theoria interpolationis, Werke* **3**, 265. Encke; *Über Interpolation, Berl. astron. Jahrbuch* 1830; *Über mechanische Quadratur*, ebenda 1837, 1862; *Ges. Abh.* **1**, 1, 21, 61. Cauchy, *Mém. sur l'interpolation, J. de math.* **2** (1837) 193. F. Neumann, *Astr. Nachr.* **15** (1838) 313. Villarceau, *Add. à la Connaiss. des Temps*

1852, p. 129. Genocchi, *Ann. di mat.* (1) **6** (1864). Tschebyscheff, *Petersburg. Mém.* (7) **1** (1859); *Œuvres* **1**, p. 387, und ebenda p. 539. Hermite, *C. R.* **48** (1859) 62; *J. f. Math.* **84** (1878) 70. Merrifield, *Report Brit. Assoc.* 1880. Gram, *J. f. Math.* **94** (1883) 41. Méray, *Ann. Éc. norm.* (3) **1** (1884) 165. Harzer, *Astr. Nachr.* **115** (1886) 337. Radau, *Études sur les formules d'interpolation*, Paris 1891. Netto, *Math. Ann.* **42** (1893) 453. Pincherle, *Bologna Mem.* (5) **3** (1893) 293. Tisserand, *Mécanique céleste* (1896), Vol. 4, p. 15 (Leverriers Interpolation durch periodische Reihen). Bruns, *Astr. Nachr.* **146** (1898) 161. Runge, *Ztschr. f. Math.* **48** (1903) 443. De la Vallée Poussin, *Brux. Acad. roy. Bull.* (1908) 319. Schließlich sei das kürzlich erschienene Lehrbuch von Thiele, *Interpolationsrechnung* (1910) genannt.

§ 3. Summen.

Zu der Differenz invers ist die *Summe*. Wenn die Funktion $f(x)$ gegeben ist, so sucht man die Funktion $\varphi(x)$, welche der Funktionalgleichung genügt:

$$\varphi(x+h) - \varphi(x) = f(x),$$

deren Differenz also die gegebene Funktion ist. Durch die vorstehende Gleichung wird die Funktion:

$$\varphi(x) = \sum f(x)$$

für eine Reihe von Werten x, die in dem Intervall h aufeinanderfolgen, bis auf eine willkürliche Konstante, die *Summationskonstante*, bestimmt. Für die Summen gelten analoge Gleichungen wie für die Differenzen. Es ist:

$$\sum C f(x) = C \sum f(x) \quad (C = \text{const.}),$$

$$\sum [\varphi(x) + \psi(x)] = \sum \varphi(x) + \sum \psi(x),$$

außerdem gilt die Formel der *partiellen Summation*:

$$\sum \varphi(x) \cdot \Delta\psi(x) = \varphi(x) \cdot \psi(x) - \sum \psi(x+h) \cdot \Delta\varphi(x).$$

Ferner wird, bis auf die noch hinzutretende Summationskonstante:

$$\sum x(x-h)\dots(x-(n-1)h) = \frac{1}{(n+1)h} x(x-h)\dots(x-nh).$$

Allgemein ist die Summe einer ganzen Funktion n^{ten} Grades

$f_n(x)$ eine ganze Funktion $(n+1)^{\text{ten}}$ Grades $f_{n+1}(x)$. Ist $A_0 x^n$ das höchste Glied in $f_n(x)$, so beginnt $f_{n+1}(x)$ mit $A_0 \frac{x^{n+1}}{n+1}$.

Endlich gelten die Formeln:

$$\sum \frac{1}{x(x+h)\dots(x+nh)} = -\frac{1}{nh}\frac{1}{x(x+h)\dots(x+(n-1)h)},$$

$$\sum a^x = \frac{1}{a^h - 1} a^x,$$

$$\sum \sin u = -\frac{\cos\left(u - \frac{h}{2}\right)}{2\sin\frac{h}{2}}, \qquad \sum \cos u = \frac{\sin\left(u - \frac{h}{2}\right)}{2\sin\frac{h}{2}}.$$

Von den unbestimmten Summen, die noch eine willkürliche Konstante enthalten, geht man zu den bestimmten Summen über, indem man die Konstante dadurch festlegt, daß für einen bestimmten Wert des Argumentes die Summe verschwinden soll. Dieser Wert x_0 heißt die untere Grenze der bestimmten Summe, und man schreibt diese in der Form:

$$\sum_{x_0}^{x} f(x).$$

Aus der vorangestellten Definition des Begriffes der Summen ergibt sich dann sofort, daß:

$$\sum_{x_0}^{x_0 + nh} f(x) = f(x_0) + f(x_0 + h) + \dots + f(x_0 + (n-1)h)$$

ist. Nimmt man z. B.:

$$f(x) = x,$$

so erhält man:

$$\sum_0^x x = \frac{x(x-h)}{2h}$$ (Summenformel der arithmetischen Reihe);

nimmt man:

$$f(x) = a^x,$$

so wird:

$$\sum_0^x a^x = \frac{a^x - 1}{a^h - 1}$$ (Summenformel der geometrischen Reihe).

Für $f(x) = x^3$ ergibt sich die Summe:

$$\sum_0^x x^3 = \frac{1}{h}\left[\frac{x(x-h)}{2}\right]^2 = \frac{1}{h}\left[\sum_0^x x\right]^2.$$

Der wesentliche Inhalt dieser Formel war nach M. Cantor

(*Gesch. d. Math.* **1**, 724) schon im 11. Jahrhundert den Arabern bekannt.

Man findet ferner:

$$\sum_h^u \cos u = \frac{\sin\left(u - \frac{h}{2}\right) - \sin\frac{h}{2}}{2\sin\frac{h}{2}},$$

$$\sum_h^u \sin u = \frac{\cos\frac{h}{2} - \cos\left(u - \frac{h}{2}\right)}{2\sin\frac{h}{2}}.$$

Wir wollen nun die Summen

$$\varphi_{n+1}(x) = \sum_0^x \frac{x^n}{n!}$$

betrachten, die nach J. L. Raabe als *Bernoullische Funktionen* bezeichnet werden (*Die Jacob-Bernoullische Funktion.* Zürich 1848). Nach dem, was wir oben über die Summe einer ganzen Funktion n^{ten} Grades gesagt haben, können wir schreiben:

$$\varphi_{n+1}(x) = \frac{x^{n+1}}{(n+1)!} + A_1 h \frac{x^n}{n!} + A_2 h^2 \frac{x^{n-1}}{(n-1)!} + \cdots + A_n h^n x.$$

Da:

$$\varphi_{n+2}(x) = \int_0^x \varphi_{n+1}(x)\, dx,$$

ändern sich, wenn man von $\varphi_{n+1}(x)$ zu $\varphi_{n+2}(x)$ übergeht, die Konstanten $A_1 \ldots A_n$ nicht, es tritt nur eine neue Konstante A_{n+1} hinzu. Für $x = h$ wird $\varphi_{n+1}(x) = 0$; somit ergeben sich zur Bestimmung der Koeffizienten die Gleichungen:

$$\frac{1}{2!} + A_1 = 0,$$

$$\frac{1}{3!} + \frac{A_1}{2!} + A_2 = 0,$$

$$\frac{1}{4!} + \frac{A_1}{3!} + \frac{A_2}{2!} + A_3 = 0$$

. . . . . .

So findet man: $A_1 = -\frac{1}{2}$, $A_2 = \frac{1}{12}$, $A_3 = 0$, $A_4 = -\frac{1}{720}$, $A_5 = 0$, $A_6 = \frac{1}{30240}$ usw. *Jedes A mit ungeradem Index, ausgenommen* A_1, *verschwindet.* Die Funktionen $\varphi_{2k}(x)$ mit

geradem Index enden also mit dem Glied, das x^2 enthält. Ferner wird:

$$\varphi_{2k}(h-x) = \varphi_{2k}(x);$$

dagegen:

$$\varphi_{2k+1}(h-x) = -\varphi_{2k+1}(x), \quad \text{woraus } \varphi_{2k+1}\left(\frac{h}{2}\right) = 0.$$

Die A mit geradem Index, die außer A_1 allein übrig bleiben, sind abwechselnd positiv und negativ. Es wird gesetzt:

$$A_1 = -B_1 \quad \text{und} \quad A_k = \frac{B_k}{k!} \qquad (k=2,3,\ldots),$$

also

$$B_1 = \tfrac{1}{2}; \quad B_3 = B_5 = B_7 = \cdots = 0.$$

Für B_{2k} erhält man dann die Reihe:

$$B_{2k}^{\,!} = (-1)^{k-1}\, 2\, \frac{(2k)!}{(2\pi)^{2k}}\left[1 + \frac{1}{2^{2k}} + \frac{1}{3^{2k}} + \frac{1}{4^{2k}} + \cdots\right].$$

Diese Entwicklung hat Euler gegeben in den *Comment. Acad. Petropol.* **12** (1740) p. 73. Man vergleiche auch Schlömilch, *Ztschr. f. Math.* **1** (1856) 200, Sonin, *J. f. Math.* **116** (1896) 138. Die Zahlen B_{2k} heißen *Bernoullische Zahlen.* Sie können durch die symbolische Gleichung

$$(B+1)^n - B^n = n \qquad (n=1,2,3,\ldots)$$

definiert werden, in der nach der Ausrechnung B^k durch B_k zu ersetzen ist. Die ersten unter ihnen sind:

$$B_2 = \frac{1}{6},\ B_4 = -\frac{1}{30},\ B_6 = \frac{1}{42},\ B_8 = -\frac{1}{30},\ B_{10} = \frac{5}{66},$$

$$B_{12} = -\frac{691}{2730},\ B_{14} = \frac{7}{6},\ B_{16} = -\frac{3617}{510},\ B_{18} = \frac{43867}{798},\ B_{20} = -\frac{174611}{330}.$$

Nach von Staudt (*J. f. Math.* **21** (1840) 372) ist

$$B_{2k} = n - \sum \frac{1}{l},$$

wo n eine ganze Zahl und die Summation nach allen Primzahlen l zu nehmen ist, für die $l-1$ ein Teiler von $2k$ ist. So wird

$$B_{10} = 1 - \frac{1}{2} - \frac{1}{3} - \frac{1}{11}, \quad B_{16} = -6 - \frac{1}{2} - \frac{1}{3} - \frac{1}{5} - \frac{1}{17}.$$

Die ersten 62 von Null verschiedenen Bernoullischen Zahlen hat Adams, *J. f. Math.* **85** (1878) 269, die ersten 90 Ssere-

brennikow, *Petersb. Abh.* (8) **16**, No. 10 (1905), berechnet. Literatur über diese Zahlen bei Ely, *Amer. J. of Math.* **5** (1882) 228.

Aus der Definitionsgleichung der Summe entsteht eine *Differenzengleichung*, wenn die Differenz der gesuchten Funktion $y = \varphi(x)$ nicht unmittelbar als eine bekannte Funktion von x, sondern als Funktion von x und y zusammen gegeben ist, insbesondere als eine für y lineare Funktion:

$$\Delta y = P_x y + Q_x,$$

wo P_x, Q_x bekannte Funktionen von x bezeichnen. Schreibt man:

$$y = y_x, \qquad \Delta y = y_{x+1} - y_x,$$

indem man das Intervall der Argumente $h = 1$ annimmt, so wird die vorige Gleichung von der Form:

$$y_{x+1} - A_x y_x = Q_x.$$

Allgemein heißt jede Gleichung von der Form:

$$y_{x+m} + A_x y_{x+m-1} + A_x' y_{x+m-2} + \cdots + A_x^{(m-1)} y_x = Q_x$$

eine lineare Differenzengleichung m^{ter} Ordnung, indem man als Differenzengleichung schlechthin jede Gleichung zwischen den Werten des Argumentes x und den Funktionswerten $y_x, y_{x+1} \cdots$ für eine von x anfangende Reihe äquidistanter Argumentenwerte bezeichnet.

Näheres über Differenzengleichungen siehe in Kap. X.

§ 4. Zusammenhang zwischen Summen und Integralen. Mechanische Quadratur.

Wie wir oben von dem Zusammenhange zwischen Differenzen und Differentialquotienten gesprochen haben, so kann man auch eine Beziehung zwischen Summen und Integralen suchen.

Auf der einen Seite kann man die Summe durch ein Integral mit Zuhilfenahme sukzessiver Differentialquotienten ausdrücken. Dieser Ausdruck wird durch die Eulersche Formel geliefert:

$$h \cdot \sum_a^b f(x) = \int_a^b f(x)\,dx + A_1 h[f(b) - f(a)] + A_2 h^2[f'(b) - f'(a)] + \cdots + A_n h^n[f^{(n-1)}(b) - f^{(n-1)}a] + R_{n+1}.$$

A_1, A_2 ... bedeuten hierin die früher so bezeichneten Zahlen, und es ist $h = \frac{b-a}{n}$. Das Restglied R_{n+1} hat Poisson, *Par.*

mém. **6**, 580 (1826) hinzugefügt; es wird nach Jacobi (*J. f. Math.* **12** (1834) 263, *Werke* **6**, 64) in folgender Form dargestellt:

$$R_{n+1} = -\int_0^h \varphi_{n+1}(u)\,du \sum_x^b {}_a f^{(n+1)}(x+h-u),$$

wenn $\varphi_{n+1}(u)$ wie oben die Bernoullische Funktion bedeutet, und es wird

$$R_{n+1} = A_{n+1} h^{n+2} f^{(n+1)}(\xi) \qquad (a < \xi < b).$$

Die Formel ohne Restglied wurde von Euler gegeben, *Comment. Acad. Petrop.* **6** (1738) 68 und ebenda **8** (1741) 147. Man sehe auch Malmstén, *Acta math.* **5** (1884) 1, Sonin, *Ann. éc. norm.* (3) **6** (1889) 257, *C. R.* **108** (1889) 725, Kronecker, *Vorl. üb. die Theorie d. Integrale* (1894), 130—156.

Als erstes Beispiel nehmen wir:

$$f(x) = x^m, \qquad b = x, \quad a = 0,$$

so wird:

$$h \sum_0^x x^m = \frac{x^{m+1}}{m+1} - \tfrac{1}{2} h x^m + \tfrac{1}{12} m h^2 x^{m-1} \ldots.$$

Nimmt man zweitens:

$$f(x) = e^{ux}, \qquad b = 1, \quad a = 0, \quad h = 1,$$

so wird:

$$\frac{1}{e^x - 1} = \frac{1}{x} + A_1 + A_2 x^2 + A_4 x^4 + \cdots.$$

Nimmt man:

$$f(x) = \log x, \qquad b = m, \quad a = 1, \quad h = 1,$$

und addiert zu beiden Seiten der Gleichung $\log m$, so erhält man die Stirlingsche Formel (Stirling, *Methodus differentialis*, 1730, p. 135):

$$\log(m!) = \log\sqrt{2\pi} + \left(m + \frac{1}{2}\right)\log m - m + A_2 \frac{1}{m} + \cdots + A_{2k} \frac{(2k-2)!}{m^{2k-1}} + \theta A_{2k+2} \frac{(2k)!}{m^{2k+1}} \qquad (0 < \theta < 1).$$

Auf der anderen Seite kann man nun auch näherungsweise ein Integral durch Summen darzustellen suchen. Diese angenäherte Auswertung heißt *mechanische Quadratur*. Die einfachsten Formeln, die dies leisten, sind die *Trapezformel* und die *Simpsonsche Regel*.

Man denkt sich das Intervall von a bis b in n gleiche

Teile h zerlegt, so daß $b - a = nh$, und bezeichnet die Funktionswerte an den Grenzstellen $x_\mu = a + \mu h$ der Teile mit:

$$y_0, y_1, y_2 \cdots y_n,$$

die Funktionswerte in der Mitte der Teile durch:

$$y_{\frac{1}{2}}, y_{\frac{3}{2}} \cdots.$$

Dann wird die *Trapezformel*:

$$\int_a^b f(x)\,dx = \frac{h}{2}(y_0 + 2y_1 + 2y_2 + \cdots + 2y_{n-1} + y_n) - R_n$$

mit:

$$R_n = n\frac{h^3}{12}f''(\xi), \qquad a < \xi < b.$$

Die *Simpsonsche Regel* lautet:

$$\int_a^b f(x)\,dx = \frac{h}{6}(y_0 + 4y_{\frac{1}{2}} + 2y_1 + 4y_{1\frac{1}{2}} \cdots + 4y_{n-\frac{1}{2}} + y_n) - R_n$$

mit:

$$R_n = n\frac{h^5}{2880}f^{IV}(\xi), \qquad a < \xi < b.$$

Simpson, *Math. diss.*, London 1743, jedoch bereits bei Gregory, *Exerc. geom.*, 1668 (vgl. Heinrich, *Bibl. math.* (3) **1**, 90 (1900)) und Toricelli, *Quadrat. parabolae*, 1644 (vgl. Aubry, *J. de math. élem.* (5) **21** (1897)).

Die erste Formel, welche die angenäherte Quadratur in konsequenten Zusammenhang mit der Theorie der Interpolation brachte, war die *Cotessche Formel*. In ihr wird das zu berechnende Integral ersetzt durch das Integral einer ganzen Funktion n^{ten} Grades, welche für $x_0, x_1 \ldots x_n$ dieselben Werte $y_0, y_1 \ldots y_n$ annimmt wie die zu integrierende Funktion. Setzt man:

$$\varphi(x) = (x - x_0)(x - x_1) \cdots (x - x_n)$$

und schreibt:

$$\int_a^b f(x)\,dx = (b - a)[\underset{n}{H}{}^{0} y_0 + \underset{n}{H}{}^{1} y_1 + \cdots + \underset{n}{H}{}^{n} y_n],$$

indem man sich $f(x)$ durch die Lagrangesche Interpolationsformel bestimmt denkt, so hat man zu setzen:

$$\underset{n}{H}{}^{\mu} = \int_a^b \frac{\varphi(x)}{x - x_\mu} \cdot \frac{dx}{(b - a)\,\varphi'(x_\mu)}.$$

Substituiert man hierin:

$$x = a + hz, \qquad x_\mu = a + h\mu,$$

so kann man das Integral von der besonderen Wahl der Grenzen unabhängig machen und schreiben:

$$\overset{\mu}{H}_n = \int_0^n \frac{z(z-1)\cdots(z-\mu+1)(z-\mu-1)\cdots(z-n)}{\mu(\mu-1)\cdots 1\cdot(-1)\cdots(\mu-n)}\,dz.$$

Hierbei wird: $\overset{\mu}{H}_n = \overset{n-\mu}{H}_n, \qquad \sum_{\mu=0}^{n} \overset{\mu}{H}_n = 1.$

Wir stellen in der folgenden Tabelle die von Cotes (*Harmonia mensurarum*, 1722) berechneten Werte für $n = 1$ bis $n = 10$ zusammen:

$\overset{\mu}{H}_n = \overset{n-\mu}{H}_n$	μ					
	0	1	2	3	4	5
$n = 1$	$\frac{1}{2}$	$\frac{1}{2}$				
$n = 2$	$\frac{1}{6}$	$\frac{4}{6}$	$\frac{1}{6}$			
$n = 3$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$		
$n = 4$	$\frac{7}{90}$	$\frac{32}{90}$	$\frac{12}{90}$	$\frac{32}{90}$	$\frac{7}{90}$	
$n = 5$	$\frac{19}{288}$	$\frac{75}{288}$	$\frac{50}{288}$	$\frac{50}{288}$	$\frac{75}{280}$	$\frac{19}{288}$
$n = 6$	$\frac{41}{840}$	$\frac{216}{840}$	$\frac{27}{840}$	$\frac{272}{840}$	$\frac{27}{840}$	$\frac{216}{840}$
$n = 7$	$\frac{751}{17280}$	$\frac{3577}{17280}$	$\frac{1323}{17280}$	$\frac{2989}{17280}$	$\frac{2989}{17280}$	$\frac{1323}{17280}$
$n = 8$	$\frac{989}{28350}$	$\frac{5888}{28350}$	$-\frac{928}{28350}$	$\frac{10496}{28350}$	$-\frac{4540}{28350}$	$\frac{10496}{28350}$
$n = 9$	$\frac{2857}{89600}$	$\frac{15741}{89600}$	$\frac{1080}{89600}$	$\frac{19344}{89600}$	$\frac{5778}{89600}$	$\frac{5778}{89600}$
$n = 10$	$\frac{16067}{598752}$	$\frac{106300}{598752}$	$-\frac{48525}{598752}$	$\frac{272400}{598752}$	$-\frac{260550}{598752}$	$\frac{427368}{598752}$

Gauß hat den Gedanken verfolgt, die Werte des Argumentes $x_0, x_1 \ldots x_n$, welche die Grundlage der Interpolation bilden, nicht durch Teilung des ganzen Intervalles in gleiche Teile zu gewinnen, sondern sie so zu wählen, daß die Genauigkeit der Formel eine möglichst große wird. Dies sucht er zu erreichen, indem er die Forderung aufstellt, daß die Formel nicht bloß exakt sein soll für die Funktion n^{ten} Grades, welche für $x_0, x_1 \ldots x_n$ die vorgeschriebenen Werte annimmt, sondern für jede ganze Funktion bis zum $(2n+1)^{\text{ten}}$ Grade, welche dieser Bedingung genügt. Das Schwergewicht liegt somit in der Festlegung der Funktion:

$$\varphi(x) = (x - x_0)(x - x_1) \cdots (x - x_n),$$

welche, gleich Null gesetzt, die Argumentenwerte $x_0 \ldots x_n$ liefert Das Resultat ist, daß der aufgestellten Forderung genügt wird, wenn:

$$\varphi(x) = \frac{d^{n+1}}{dx^{n+1}} [(x-a)^{n+1}(x-b)^{n+1}]$$

angenommen wird. In der so entstehenden Gleichung $\varphi(x) = 0$ werden die von Fall zu Fall wechselnden Grenzen a, b zu den festen Werten 0,1, wenn

$$x = a + (b - a)t$$

substituiert wird. Dann wird die Gleichung:

$$\frac{d^{n+1}}{dt^{n+1}} [t^{n+1}(1-t)^{n+1}] = 0,$$

und ihre Wurzeln können ein für allemal berechnet werden. Diese Berechnung hat Gauß bis auf 16 Dezimalen durchgeführt. Die folgende Tabelle gibt die Wurzeln für die ersten 7 Grade auf 6 Dezimalen:

Grad $n+1$	Wurzeln t						
	t_0	t_1	t_2	t_3	t_4	t_5	t_6
1	0,5						
2	0,211325	0,788675					
3	0,112702	0,5	0,887298				
4	0,069432	0,330009	0,669991	0,930568			
5	0,046910	0,230765	0,5	0,769235	0,953090		
6	0,033765	0,169395	0,380690	0,619310	0,830605	0,966235	
7	0,025446	0,129234	0,297077	0,5	0,702923	0,870766	0,974554

Es sind nun weiter zu berechnen die Integrale:

$$E_\mu = \int_0^1 \frac{(t-t_0)\dots(t-t_{\mu-1})\,(t-t_{\mu+1})\dots(t-t_n)}{(t_\mu-t_0)\dots(t_\mu-t_{\mu-1})\,(t_\mu-t_{\mu+1})\dots(t_\mu-t_n)}\,dt.$$

Die Werte derselben für die ersten 7 Grade sind in der folgenden Tabelle zusammengestellt:

Grad $n+1$	E_0	E_1	E_2	E_3	E_4	E_5	E_6
1	1						
2	$\frac{1}{2}$	$\frac{1}{2}$					
3	$\frac{5}{18}$	$\frac{8}{18}$	$\frac{5}{18}$				
4	0,173927	0,326073	0,326073	0,173927			
5	0,118463	0,239314	0,284444	0,239314	0,118463		
6	0,085662	0,180381	0,233957	0,233957	0,180381	0,085662	
7	0,064742	0,139853	0,190915	0,208980	0,190915	0,139853	0,064742

Die Gaußsche Formel schreibt sich dann in der einfachen Gestalt:

$$\int_a^b f(x)\,dx = (b-a)[E_0 f(a+(b-a)t_0) + \dots + E_n f(a+(b-a)t_n)].$$

Außer der Arbeit von Gauß (*Comment. Gotting.* **3** (1814/15), Selbstanzeige *Gött. gel. Anz.* 1814, *Werke* **3**, 163, 202) vergleiche man u. a.: Jacobi, *J. f. Math.* **1** (1826) 301, *Werke* **4**, 3. Christoffel, *J. f. Math.* **55** (1858) 61. Tschebyscheff, *Journ. de Math.* (2) **19** (1874) 19. Mansion, *Mathesis* **1** (1881), Supplem., p. 1; *Bruxelles Ann. Soc. Scient.* **5** B (1881) 231; *Bruxelles Bullet.* (3) **11** (1886) 293; *Comptes rendus* **102** (1886) 412 (*Bestimmung des Restes in der Gaußschen Formel*). Stieltjes, *Ann. Éc. norm.* (3) **1** (1884) 409; *Comptes rendus* **99** (1884) 850. Markoff, *Math. Ann.* **25** (1885) 417. Man vgl. auch den Bericht von Burkhardt, Math.-Ver. **10** 2, § 88.

Lehrbücher der Differenzenrechnung:

S. F. Lacroix, *Traité des différences,* Paris 1800, 3. Bd. des *Traité du calcul differentiel etc.* J. F. W. Herschel, *Collection of examples of the applications of the calculus of finite diffe-*

rences, Cambridge 1820. O. Schlömilch, *Theorie der Differenzen und Summen,* Halle 1848. G. Boole, *A Treatise on the calculus of finite differences,* Cambridge 1860, deutsch von Schnuse, Braunschweig 1867 u. d. T.: *Die Grundlehren der endlichen Differenzen- und Summenrechnung.* A. Markoff, *Differenzenrechnung,* deutsch von Friesendorff und Prümm, Leipzig 1896. E. Pascal, *Calcolo delle differenze finite,* Milano 1897. H. Bruns, *Grundlinien des wissenschaftl. Rechnens,* Leipzig 1903. O. Biermann, *Vorlesungen über mathematische Näherungsmethoden,* Braunschweig 1905. D. Seliwanoff, *Differenzenrechnung, Teubners Sammlung* Bd. 13, Leipzig 1904.

Verbesserungen.

S. 147, die letzten vier Zeilen sind zu streichen.
„ 213, Zeile 22 v. o. lies: $|z, \alpha+\beta z|$ statt $|z, \alpha z+\beta|$.

Druck von B. G. Teubner in Leipzig

QA
37
P375
1910
Bd.1
T.1

Pascal, Ernesto
Repertorium der höheren Mathematik 2. völlig umgearb. Aufl.

Physical &
Applied Sci.

PLEASE DO NOT REMOVE
CARDS OR SLIPS FROM THIS POCKET

UNIVERSITY OF TORONTO LIBRARY